

**The Genetic, Cellular, and Evolutionary Basis of Skin Coloration in the Highly
Polymorphic Poison Frog, *Oophaga pumilio***

by

Layla Renee Freeborn

B.S. Biology, Utica College of Syracuse University, 2008

M.S. Biology, Southeastern Louisiana University, 2010

Submitted to the Graduate Faculty of the
Dietrich School of Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DIETRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Layla Renee Freeborn

It was defended on

October 5, 2020

and approved by

Dr. Miler T. Lee, Assistant Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Mark Rebeiz, Associate Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Martin Turcotte, Assistant Professor,
Department of Biological Sciences, University of Pittsburgh

Dr. Marcus R. Kronforst, Professor,
Department of Ecology & Evolution, University of Chicago

Dr. Nathan Morehouse, Associate Professor,
Department of Biological Sciences, University of Cincinnati

Dissertation Advisor: Dr. Corinne L. Richards-Zawacki, Associate Professor,
Department of Biological Sciences, University of Pittsburgh

Copyright © by Layla Renee Freeborn

2020

**The Genetic, Cellular, and Evolutionary Basis of Coloration in the Highly Polymorphic
Poison Frog, *Oophaga pumilio***

Layla Renee Freeborn, Ph.D.

University of Pittsburgh, 2020

Studies of phenotypic variation are undisputedly important to understanding the processes of evolution. Animal coloration is important in many ecological contexts, including thermoregulation, predator avoidance, male-male competition, and mate choice. Given its varied function, animal coloration is frequently characterized by phenotypic variation among and within species. The Central American strawberry poison frog, *Oophaga pumilio*, is known for exhibiting striking color polymorphism across the Bocas del Toro archipelago of Panama. Despite many studies that aim to understand the selective pressures that generated or maintain this variation, relatively little is known about its underlying genetic and morphological basis. Furthermore, the evolution of color variation in this species has been addressed by phylogenetics and population genetics studies, but they have been limited by the challenges associated with studying recently separated populations. To bridge this gap in our knowledge, I investigated the proximate mechanisms of color variation by leveraging the power of next-generation sequencing to find loci associated with color differences. I found that the genes involved in variation in dorsal background coloration are located across broad regions of the genome, rather than being co-localized in one or a few nearby regions. Next, I investigated the proximate mechanism of color variation by examining the presence and absence of pigment types, ratios of pigments, skin components, and chemical compositions of carotenoid-containing cells. My results suggest that large color

variations in *O. pumilio* are explained by differences in pigment proportions and carotenoid suites, not by the absence of pigment cell types. Finally, I applied a multi-species coalescent approach to a SNP dataset derived from the extensive sampling of *O. pumilio* from Bocas del Toro, then perform ancestral character state reconstructions of dorsal coloration. These analyses provide a high-resolution estimate of relationships for *O. pumilio*, despite the challenges associated with recent timescales. The results suggest that vicariance explains pattern of genetic differentiation and support the convergent evolution of dull dorsal coloration. Overall, this dissertation provides an important first step toward understanding the proximate mechanisms of phenotypic variation and provides a much-needed phylogenetic backbone for future studies of this system.

Table of Contents

Preface.....	xiv
1.0 A Population Genomic Scan and Genetic Linkage Map for the Highly Polymorphic Poison Frog, <i>Oophaga pumilio</i>, Identifies Linked Candidate Coloration Markers	1
1.1 Summary	1
1.2 Introduction	2
1.3 Methods	7
1.3.1 Sampling Strategy: Polymorphic Populations	7
1.3.2 Tissue Collection from Wild Populations	8
1.3.3 Sampling Strategy: Lab Colony.....	9
1.3.4 DNA Extraction and Library Preparation	11
1.3.5 Quality Filtering of Raw Reads	13
1.3.6 Population Genomics and Association Study	13
1.3.7 Amplifying Candidate Loci Regions in Pedigree Frogs	15
1.3.8 Mapping Candidate Loci.....	16
1.4 Results.....	17
1.4.1 Population Genomic Scan	17
1.4.2 Linkage Mapping and Mapping Candidate Coloration Loci	19
1.5 Discussion	22
2.0 The Proximate Mechanisms of Skin Color Variation in a Highly Polymorphic Poison Frog, <i>Oophaga pumilio</i>	28

2.1 Summary	28
2.2 Introduction	29
2.3 Methods	35
2.3.1 Sampling Strategy for Reflectance Spectroscopy and Morphological Analysis	35
2.3.2 Reflectance Spectroscopy and Color Assignment	37
2.3.3 Dorsal Skin Morphology	39
2.3.4 High-performance Liquid Chromatography (HPLC) and Ultraviolet-visible (UV/VIS) Spectroscopy of Carotenoid Pigments	41
2.3.5 Compositional Treatment of Proportional Pigment Data	42
2.3.6 Statistical Analysis of Morphological and Chemical Data	43
2.4 Results.....	46
2.4.1 Characterizing Coloration from Spectral Reflectance Data	46
2.4.2 Morphological and Chemical Determinants of Skin Coloration	47
2.4.2.1 The absence of pigment types from individuals and morphs	47
2.4.2.2 Total pigment and pigment proportion effects on chroma, total brightness.....	48
2.4.2.3 Dermis, epidermis, and collagen layer heights.....	51
2.4.2.4 Pigment proportions and layer heights as predictors of color groups	53
2.4.2.5 Chemical determinants and differences among color groups	56
2.5 Discussion	61

3.0 Estimating the Phylogenetic Relationships Among Closely Related Populations of	
<i>Oophaga pumilio</i> from Bocas del Toro, Panama	67
3.1 Summary	67
3.2 Introduction	68
3.3 Methods	72
3.3.1 Sampling Details.....	72
3.3.2 ddRADseq: Library Preparation, Sequencing, and Bioinformatic Processing	
.....	74
3.3.3 Population Differentiation.....	74
3.3.4 Phylogenetic Analysis	75
3.3.5 Coloration Data from Reflectance Spectroscopy	76
3.3.6 Tests of Phylogenetic Signal and Ancestral Character State Reconstructions	
of Skin Coloration	76
3.4 Results.....	78
3.4.1 Population Differentiation.....	78
3.4.2 Phylogenetic Analysis	80
3.4.3 Tests of Phylogentic Signal and Ancestral Character State Reconstructions	
.....	82
3.5 Discussion	83
Appendix A Supplement to Chapter 1	89
Appendix B Supplement to Chapter 2	91
Appendix C Supplement to Chapter 3	97
Bibliography	104

List of Tables

Table 1. The parent populations, the number of offspring contributing to the linkage map, and the number of offspring included in the genotyping of candidate coloration loci.	10
Table 2. Candidate locus ID, number of SNPs and offspring supporting each locus, and associated BLAST results for 18 candidate coloration loci.	19
Table 3. Summary of results of linear mixed effects models with pivot coordinates of pigment proportions as the predictor variables and color metrics as the explanatory variables.	49
Table 4. Coefficients of linear discriminants for LDA with and without height data (with/without).	54
Table 5. Prediction accuracy for resubstitution and leave-one-out cross validation methods.	54
Table 6. A summary of significant SIMPER results.	58
Table 7. Summary of MANTEL tests to estimate the correlation between carotenoid suites and individual color metrics.	59
Table 8. Summary of linear models of the relationship between individual carotenoid pigments and colorimetric variable.	60
Appendix Table 1. Marker segregation types used in the construction of linkage maps and the mapping of candidate loci.	89
Appendix Table 2. Forward and reverse primer sequences, annealing temperatures, and the approximate amplified fragment length for five candidate loci.	89

Appendix Table 3. Candidate regions for markers heterozygous in one or more parents..	90
Appendix Table 4. Sampling locations, abbreviations, and GPS coordinates included in the histological and HPLC datasets.....	91
Appendix Table 5. Results of linear mixed effects models with total pigment as the predictor variable and color metric as the explanatory variables.	91
Appendix Table 6. Summary of similarity percentages breakdown (SIMPER) test.....	92
Appendix Table 7. Detailed summary of linear models of the relationship between individual carotenoid pigments and color metric variable.....	92
Appendix Table 8. Sampling locations, abbreviations, and GPS coordinates included in the phylogenetics dataset.	97

List of Figures

Figure 1. Locations and photographs showing <i>O. pumilio</i> morphs present at focal populations in the Bocas del Toro region of Western Panama.....	5
Figure 2. Scatterplots of p values plotted against F_{ST} values for Dolphin Bay Preserve (top) and Cemetery Hill, Isla Bastimentos (bottom).	18
Figure 3. Sex-specific linkage map for male <i>O. pumilio</i> at LOD=2.4.	21
Figure 4. Sex-specific linkage groups for female <i>O. pumilio</i> at LOD=2.4.....	22
Figure 5. Stained histosections identifying major features of <i>O. pumilio</i> skin.....	31
Figure 6. Sampling strategy for histological and chemical analyses.....	36
Figure 7. A reflectance plot illustrating the color metrics used in this analysis.....	38
Figure 8. Visualization of dorsal (A) and ventral (B) clusters based on model-based cluster analyses.	47
Figure 9. Scatterplot of pivot coordinates of pigment proportions versus total brightness (B_1).	48
Figure 10. Scatterplot of pivot coordinates of pigment proportions (from left to right: reflecting platelets, melanins, carotenoids) versus chroma metrics (from top to bottom panel: $S1_V$, $S1_B$, $S1_G$, $S1_Y$, $S1_R$)......	51
Figure 11. Boxplots showing mean (B) epidermis, (C) dermis and (D) collagen heights, by color group.....	52
Figure 12. Stacked histograms of the first discriminant function value for samples from different color groups.	55

Figure 13. NMDS plot of HPLC abundance data from dorsal <i>O. pumilio</i> skin, illustrating differences in carotenoid composition between some <i>mclust</i> -derived color groups.	57
Figure 14. Map of Bocas del Toro, Panama, showing the location of samples included in this analysis.	73
Figure 15. Coancestry matrix from fineRADstructure based on 673 haplotypes containing 2,489 SNPs from 222 <i>O. pumilio</i> individuals.	79
Figure 16. Species tree resulting from a SNAPP analysis that included unlinked, bi-allelic SNPs from 5 individuals per ingroup OTU.	81
Figure 17. Ancestral character state reconstructions of dorsal coloration using a polymorphic evolution with ‘all rates different’ model and the SNAPP topology.	83
Figure 18. Predicted topologies based on vicariance models (A, B) and the observed topology (C).	87
Appendix Figure 1. Scatterplot illustrating the relationship between total brightness (B _i) and total carotenoids measured from HPLC.	95
Appendix Figure 2. Scatterplot illustrating the relationship between the proportion of carotenoids measured from histosections and the total carotenoids measured from HPLC.	96
Appendix Figure 3. Schematic detailing the bioinformatic steps taken to filter SNP variants.	98
Appendix Figure 4. Boxplot showing the proportion of missing alleles by collecting site.	99
Appendix Figure 5. Neighbor joining tree with a Kimura 2 substitution model, demonstrating reciprocal monophyly among the majority of OTUs identified in the coancestry matrix (Fig. 15).	100

Appendix Figure 6. Densitree resulting from a SNAPP analysis that included unlinked, bi-allelic SNPs from 5 individuals per ingroup OTU.	101
Appendix Figure 7. Species tree resulting from the SVDQuartets analysis implemented in PAUP and bootstrapped with 100 replicates.	102
Appendix Figure 8. Ancestral character state reconstructions of dorsal coloration using a polymorphic evolution with ‘all rates different’ model and the SVDQuartets topology.	103

Preface

Many thanks are in order. To my advisor, Dr. Cori Richards-Zawacki, who always set a standard for professionalism, hard work, and making lemonade out of lemons. My thesis committee: Drs. Marcus Kronforst, Nate Morehouse, Miler Lee, Mark Rebeiz, and Martin Turcotte, for their encouragement and patience. Thank you to my funding sources: the Smithsonian Tropical Research Institute, the Society for Systematic Biologists, EDEN Evo-Devo-Eco Network, Roger Thayer Stone Center for L.A. Studies at Tulane University, and Louisiana EPSCoR. My labmates- Sebastian Echeverri, Kim Howell, Michel Ohmer, Veronica Saenz, Julia Sonn, Justin Yeager, and Yusan Yang. I especially appreciate Yusan and Sebastian's eleventh hour help when they had no obligation to do so. This work would not be possible without my field assistants, particularly Julia Berkey, Christina Buelow, Garrett Freeborn (brother), Henry Bart, and Daniel Lenger.

My data collection and analyses depended on the generosity of people from numerous centers, labs, and homes across the Americas: the staff at the Smithsonian Tropical Research Institute in Bocas del Toro, Panama and La Selva Biological Station in Costa Rica, Dr. Sean Mullen's lab at Boston University, the Berkey family for their hospitality, Dr. Marcus Kronforst and Dr. Wei Zhang lab at the University of Chicago, the Center for Biological Imaging (CBI) at the University of Pittsburgh, and the National Center for Genome Analysis Support (NCGAS) at Indiana University-Bloomington. I am especially grateful to CBI's Jonathan Franks, Mara Sullivan, and Ming Sun. Dr. Sheri Sanders from NCGAS was indispensable to my phylogeographic analyses and professional development. There are countless expats and Panamanians to who let me wander around their properties looking for frogs.

Thank you to my support network, made up of dear friends, family, and Paula Kauffman Oberly. To my husband, Thomas Beach, who always reminded me that “work is just work”. To Monique Dupas Blossom, whose friendship, good humor, and unwavering support got me through the hardest of times. To my mom, whose unconditional love is a welcomed constant in my life.

Finally, to the fur babies on my lap and at my feet: Rusty, Bitsy, Juniper, and Reggie. Reginald, you’re an absolute champ, and I couldn’t have done this without you. I love you, bud.

1.0 A Population Genomic Scan and Genetic Linkage Map for the Highly Polymorphic Poison Frog, *Oophaga pumilio*, Identifies Linked Candidate Coloration Markers

1.1 Summary

Aposematic, or warning, coloration is a focal component of many well-studied animal systems and provides some of the most intriguing examples of adaptive evolution. Dramatic variation in dorsal background coloration, separation of morphs on islands, and a geologically young landscape make this species a fascinating system for studying the evolutionary processes leading to intraspecific divergence in aposematic coloration. A complete understanding of the factors leading to phenotypic divergence, in general, requires information from multiple levels of biological organization, from the genetic mechanisms underlying traits of interest to their ecological function. With this in mind, my primary aim was to take a first but necessary step toward understanding the genetic basis of phenotypic divergence in dorsal background coloration among *O. pumilio* morphs. First, I surveyed genomic variation in wild-caught frogs from polymorphic populations and identified SNPs significantly associated with dorsal background coloration. Next, I genotyped five candidate coloration regions in parents and offspring with known pedigrees, using Sanger sequences including and surrounding the candidate coloration loci. Then I generated sex-specific linkage maps for *O. pumilio* using mappable RAD markers derived from the same pedigreed individuals. I incorporated candidate coloration regions with the linkage maps to determine whether this extraordinary example of phenotypic variation implicates one or multiple genomic regions. My results suggest that the genes involved in variation in aposematic dorsal background coloration are located across broad regions of the genome, rather than being

co-localized in one or a few nearby regions. This analysis provides an exciting first step toward understanding the genetic basis of dorsal background coloration in the strawberry poison frog and sets the stage for future investigations on the evolution of phenotypic divergence.

1.2 Introduction

Understanding the evolutionary processes that lead to phenotypic diversity, and ultimately speciation, has long been a central goal for biologists. Elucidating the molecular features that collectively describe the genetic architecture of a phenotype- the number and positions of underlying genes, their relative effect sizes, and interactions among them- has been pivotal to learning about evolution (Hoban et al., 2016; Feder et al., 2012). Population genomics has emerged as a reliable approach to identify the genetic basis of divergent phenotypes in a diversity of non-model systems (Lee et al., 2014). An important outcome of this research is the ongoing development of a taxonomically broader understanding of how evolution proceeds.

Since the application of population genomics techniques to model and non-model organisms alike, molecular genetic experiments have illustrated ways that underlying genetic architecture can facilitate rapid or repeated phenotypic changes, even in the face of ongoing gene flow (Feder et al., 2012). Phenotypic traits controlled by loci of major effect can diverge more readily, due to higher per-locus selection coefficients, than if controlled by numerous loci of small effect (Nosil et al., 2009; Yeaman and Whitlock, 2011; Comeault et al., 2015). Genes of major effect have been attributed to a variety of discrete, divergent phenotypes in numerous plant and animal systems, including flower color in monkeyflowers (Bradshaw & Schemske, 2003), coat

color in mice (Hoekstra et al, 2006; Steiner et al, 2007), skeletal changes in sticklebacks (Chan et al., 2010), and dorsal coloration in lizards (Rosenblum et al., 2010).

The relative locations of genes also determine the influence of evolutionary processes, particularly those contributing to the maintenance of divergent adaptive phenotypes. For example, co-localized clusters of tightly linked functional genetic elements, referred to as ‘supergenes’, have been implicated in morphological, physiological, and behavioral phenotypes in a variety of taxa (Schwander et al., 2014). The physical proximity of loci or mechanisms that suppress recombination, most commonly via structural changes such as inversions, ensure that supergene loci are inherited together. The resulting simple Mendelian behavior allows polymorphisms in a whole suite of co-adapted traits to be maintained in a population (Tuttle et al., 2016; Zhang et al., 2017). In this way, supergenes provide an intriguing example of how the underlying genetic architecture of a trait allows for selection to maintain divergent phenotypes within a population.

Despite the popularity and tractability of phenotypes with supergene or major effect architecture, the number of loci (i.e., few genes of large effect or many genes of small effect) commonly involved in adaptation is still an unresolved question in evolutionary biology. Theoretical debate aside, a common explanation for the preponderance of major effect loci is discovery bias toward the identification of large-effect loci, particularly in quantitative trait loci (QTL) mapping (Rockman, 2012). However, the accessibility of next-generation sequencing (NGS) has allowed mapping studies with increased statistical power; genome wide association studies (GWAS) and outlier analyses have emerged as complementary approaches to traditional genetic mapping techniques, allowing mapping from wild populations or in the absence of a reference genome (Stapley et al., 2010). Nevertheless, the identification of small-effect loci as the

basis for adaptive traits, especially in non-model organisms, has largely evaded biologists (but see Santure et al., 2015 for notable exceptions).

Dorsal background coloration in the strawberry poison frog, *Oophaga pumilio*, provides a fascinating example of rapid phenotypic divergence in an adaptive trait. Throughout most of the species' range, from Nicaragua to Panama, *O. pumilio* populations are composed of individuals with red dorsal background coloration and blue or black limbs. However, populations across the Bocas del Toro archipelago and on the adjacent mainland of northwestern Panama exhibit remarkably variable dorsal coloration, including orange, blue, red, and green morphs. Although dorsal background coloration varies greatly among islands, individuals from the same island generally appear similar. Two noteworthy exceptions are the polymorphic populations on Isla Bastimentos and the Aguacate Peninsula (Fig. 1). Given the recency of island formation (1000-10000 ya), divergence in coloration from the ancestral red morph with blue limbs to the variety observed across Bocas del Toro appears to have been both recent and rapid (Gehara et al., 2013).

Coloration in *O. pumilio* has been studied in a range of contexts, including the evolution and maintenance of aposematic (warning) signals (Crothers et al., 2016), assortative mate preferences as a driver of reproductive isolation (Yang et al., 2016), and the role of natural selection, sexual selection, and genetic drift in driving and maintaining phenotypic variation (Gehara et al., 2013 and citations within). Nevertheless, relatively little is known about the genetic basis of color polymorphism in this system or for dendrobatid frogs in general. In *O. pumilio*, a wild pedigree estimated from microsatellite markers suggested that dorsal background coloration in a yellow/red polymorphic population is controlled by a single locus with two alleles and complete dominance of red over yellow (Richards-Zawacki et al., 2012). Observations from

captive breeding of these morphs, where yellow pairs always produce yellow offspring but pairs with at least one red parent produce a mix of yellow and red offspring, agree with the single locus

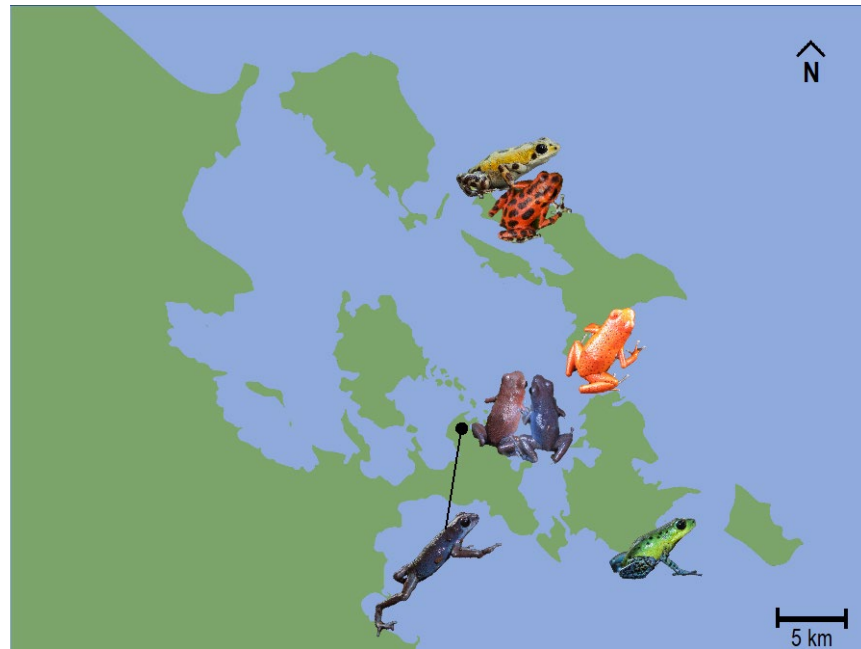


Figure 1. Locations and photographs showing *O. pumilio* morphs present at focal populations in the Bocas del Toro region of Western Panama.

Samples from the polymorphic populations located at Cemetery Hill, Isla Bastimentos (red, yellow frogs) and Dolphin Bay Preserve on the Aguacate Peninsula (blue, red frogs) were included in the association study. The lab colony, used to derive the linkage map and candidate region datasets, originated from three other populations: Punta Laurel, Isla Popa (green frog), Tranquilo Bay, Isla Bastimentos (orange frog), and Shark Hole on the Aguacate Peninsula (blue frog).

with red dominance hypothesis. However, cross-breeding between a wide range of morphs from monomorphic populations produce offspring that show a mixture of parental colors, implicating either incomplete dominance at a single locus or polygenic inheritance in the control of background dorsal coloration (Summers et al., 2004). In other dart frog systems, work on the

genetic underpinnings of phenotypic variation has focused on differences in melanistic patterning (Vestergaard et al, 2015; Posso-Terranova & Andrés, 2017), with fewer focusing on large shifts in dorsal background coloration (Twomey et al., 2020). Until recently, large genome sizes, long generation times, and lack of genetic resources for dendrobatids have prevented genetic mapping or candidate gene approaches from elucidating the underlying genomic architecture of non-melanistic color polymorphism.

Here I examine the underlying genetic architecture of color variation in the highly polymorphic *O. pumilio* populations from Bocas del Toro, Panama using samples collected from naturally polymorphic populations and animals of known pedigree from the laboratory colony at the University of Pittsburgh. I leveraged two polymorphic populations from the Bocas del Toro archipelago to study the genetic basis of dorsal background coloration. Breeding among morphs in these polymorphic populations, as evidenced by a pedigree analysis (Richards-Zawacki et al., 2012) and lack of genetic structure (between morph $F_{ST} \sim 0$, Yeager et al., *in review*), implies that the only genetic variants among sympatric color morphs should be at loci located in proximity to color genes. Furthermore, the presence of multiple polymorphic populations means that false-positive color-associated markers can be excluded by focusing on markers that are common to both populations. The laboratory colony consists of wild-caught individuals and several generations of frogs with known pedigrees, the latter allowing the construction of a genetic linkage map. The rise of reduced-representation sequencing approaches has coincided with the application of restriction-site-associated DNA (RAD) sequencing to linkage mapping studies for many non-model organisms (Puritz et al., 2013). RADseq takes advantage of the random, ubiquitous occurrence of restriction sites to generate and then sequence thousands of markers from throughout the genome, resulting in many more parent and offspring genotypes than previously possible.

In this study I aim to address the underlying genetic architecture of phenotypic differentiation in background dorsal coloration for the highly polymorphic poison frog *O. pumilio* by using a powerful combination of population genomics and genetic mapping. First, I conducted population genomic scans of two polymorphic populations to identify genetic loci associated with differences in coloration. Next, I genotyped areas including and/or surrounding these candidate coloration loci in parents and offspring with known pedigrees. Finally, I generated a genetic linkage map from these pedigreed individuals and mapped the candidate coloration loci to ask whether one or multiple genomic regions are implicated in this phenotypic variation. My overarching goal was to provide an important first step towards understanding the genetic architecture of this extraordinary example of phenotypic variation. The number of genes and their proximity to one another, two features of genetic architecture, can determine the propensity for rapid phenotypic evolution. Therefore, teasing apart the genetic basis of coloration in *O. pumilio* is essential to understanding the evolutionary forces shaping this adaptive phenotypic trait.

1.3 Methods

1.3.1 Sampling Strategy: Polymorphic Populations

The polymorphic populations sampled for this study have been previously described by several authors (Richards-Zawacki et al., 2012; Dugas et al., 2015, Yang et al., 2016). One population, referred to here as Cemetery Hill, is located on the northwest side of Isla Bastimentos (9.3458°N, 82.206°W; Fig. 1) in a 0.75 ha area bordered on one side by a cemetery. Frogs in this population are characterized by black or brown dorsal patterning that varies greatly among

individuals in terms of size and number of spots. Background dorsal coloration is either red or yellow, resulting in two distinct morphs. Ventral skin, regardless of dorsal coloration, lacks patterning and appears white. A visual inspection of photos taken during a mark-recapture survey estimates the frequency of yellow morphs to be 0.35; this same study reported essentially no genetic differentiation between color morphs ($F_{ST} = 0.003$, $P=0.17$ based on data from 11 microsatellite loci; Richards-Zawacki et al., 2012).

A second population, referred to here as Dolphin Bay Preserve, is located on the northern tip of the Aguacate Peninsula of mainland Panama (9.2210°N, 82.2185°W; Fig. 1) in a ~40 ha area. Frogs in this population lack notable dorsal patterning; background dorsal coloration appears to be a continuum from blue to red, with most frogs falling in the intermediate (brown) range. Based on a mark-recapture survey of the area, the frequencies of red, blue, and intermediate frogs have been estimated to be 0.06, 0.22, and 0.72, respectively (Dugas, unpublished data). Ventral coloration varies greatly among individuals, ranging from all blue to a mix of red and blue. A STRUCTURE analysis of variation in microsatellite markers sampled from a region including the Dolphin Bay Preserve found no evidence for genetic differentiation among morphs in this population (Yeager et al., *in review*).

1.3.2 Tissue Collection from Wild Populations

I captured by hand 15 adult individuals per morph from the polymorphic populations described above (n=60). To avoid individuals with intermediate coloration, I selected frogs that represented the purest form of each color morph (e.g., the bluest blues, the reddest reds, and the most yellow yellows). I transported frogs in plastic terraria (approximately 37cm x 22cm x 25cm) from the site of collection to a laboratory at the Smithsonian Tropical Research Institute (STRI)

on Isla Colón, Bocas del Toro, Panama, where they were housed for no more than 12h before processing. During and after transport each terrarium contained water and moist substrate and held no more than 10 frogs. I photographed the dorsal and ventral sides of each frog against a photographic gray card using a Canon PowerShot ELPH 100 HS and measured snout-to-vent length (SVL) with a dial caliper. I euthanized frogs by the application of a 7% topical benzocaine gel to the ventral skin, following an IACUC-approved protocol (Tulane University No. 0414, STRI No. 071207) and with permission from the Panamanian government (MiAmbiente; Permit Nos. SC/A-34-11 and SEX/A-73-11). I dissected toe, liver, and thigh muscle tissues for genetic analysis and preserved samples in either a room-temperature tissue preservative (salt-saturated dimethyl sulfoxide and EDTA solution) or 95% ethanol.

1.3.3 Sampling Strategy: Lab Colony

The laboratory colony includes wild-caught frogs (P), their “cross-bred” or “full bred” offspring (F₁), and a back-crossed (F₂) generation. The wild-caught frogs were collected from three populations in Bocas del Toro, Panama (Fig. 1): Punta Laurel, Isla Popa (consisting of frogs with green dorsal background coloration; 9°8'28.90"N, 82° 7'34.37"W), Tranquilo Bay, Isla Bastimentos (red frogs; 9°15'16.90"N, 82° 8'47.90"W), and Shark Hole on the Aguacate Peninsula (blue frogs; 9°12'47.12"N, 82° 12'49.28"W).

At the start of this study, the colony was limited to multiple small, two-generation families with one to 10 F₁ hybrid offspring each but no F₂ generation, and multiple small, three-generation families with one to four F₂ back-crossed offspring. Although F₂ families have the potential to provide valuable information about color pattern segregation, I excluded these three-generation pedigrees because the small numbers of F₂ offspring, thereby limiting my analysis to the inference

of the genomic region(s) containing color-associated RAD markers. Of the eight families I selected for this mapping analysis, five are crosses between a red male from Isla Bastimentos and a green female from Isla Popa and three are crosses between a red male from Isla Bastimentos and a blue female from Shark Hole. I first selected the largest two-generation families, at the time containing between eight and 10 F₁ offspring, then included sequentially smaller families, up to a total of 50 F₁ hybrid offspring. The parents of these eight families remained paired after RAD library preparation and next-generation sequencing, ultimately expanding family sizes to include 114 F₁ hybrid offspring; this larger pedigree dataset was used for the Sanger sequencing of color-associated loci, described below. Table 1 lists, for each of the eight families, the parental origins, the number of F₁ offspring contributing to RAD-based mapping, and the number of offspring included in the Sanger sequencing of candidate coloration loci. I collected toe-clips or tadpole tails from laboratory frogs using a sterilized razor blade, according to IACUC-approved protocols (Tulane University Nos. 0382 and 0382R), and then preserved tissues in 95% ethanol.

Table 1. The parent populations, the number of offspring contributing to the linkage map, and the number of offspring included in the genotyping of candidate coloration loci.

Family Number	Male Parent Origin	Female Parent Origin	Number of F₁ Offspring, Linkage Mapping	Number of F₁ Offspring, Genotyping
1	Tranquilo Bay	Punta Laurel	10	17
2	Tranquilo Bay	Punta Laurel	8	16
3	Tranquilo Bay	Punta Laurel	10	15
4	Tranquilo Bay	Shark Hole	3	16
5	Tranquilo Bay	Shark Hole	6	15
6	Tranquilo Bay	Shark Hole	5	12
7	Tranquilo Bay	Punta Laurel	3	11
8	Tranquilo Bay	Punta Laurel	5	12

1.3.4 DNA Extraction and Library Preparation

I prepared two sets of samples for sequencing according to the single-digest RAD protocol described by Etter et al. (2011), using the restriction enzyme *SbfI* (restriction site: 5'-CCTGCAGG-3'). I prepared the first sample set, consisting of 40 wild-caught frogs (10 individuals/morph) from the polymorphic populations, at STRI's Earl S. Tupper Research and Conference Center in Panama City, Panama. First, I extracted genomic DNA in spin-columns using the Qiagen DNeasy Blood and Tissue Kit and quantified DNA using a Qubit Fluorometric Quantification Platform (Invitrogen, Carlsbad, CA). Next, I digested 1 µg of each sample with 1 µL of 20 U/µL *SbfI*-HF (New England Biolabs, Ipswich, MA), resulting in random fragments that I ligated to unique, five-digit, *SbfI*-specific barcodes. I created one large pool by combining equal volumes from each sample, then sheared three 100 µL aliquots containing ~1.3 µg DNA/aliquot for 50s with a Covaris sonicator, resulting in fragments > 1-kb. I ran twenty microliters from each sheared aliquot in separate lanes of a 1.25% agarose gel alongside a 100-bp DNA ladder (GeneRuler) for ~60m at 96V. I manually excised gel fragments spanning 300-500-bp from each lane using a sterile razorblade, then extracted DNA from gel fragments using the MinElute Gel Purification Kit. I added a second, Y-shaped adapter, with one arm complementary to the reverse amplification primer site, to the purified products. Finally, I amplified each aliquot by polymerase-chain reactions (PCR), performed with 18 cycles of amplification using RAD-amplification primers. To reduce amplification bias I conducted 10 separate PCR reactions, rather than one large-volume PCR reaction, for each aliquot. This process resulted in three amplification pools, each containing independently-amplified fragments from 40 individually-barcoded frogs. Prior to submission for sequencing, I measured the molar concentrations of each amplification pool using a Qubit (Invitrogen, Carlsbad, CA). I assessed the quality of each pool using an Agilent 2100

Bioanalyzer (Agilent, Santa Clara, CA). I sequenced two of the three pools using 100-bp paired-end sequencing on two lanes of an Illumina HiSeq2000 at the FAS Center for Systems Biology at Harvard University.

I prepared the second sample set, containing wild-caught frogs from the polymorphic populations (n=20) and lab-raised frogs from the families described above (n=66), for RAD-sequencing at Boston University. I included wild-caught frogs to replace samples from the first sequencing run that produced few high-quality reads; in some cases the same frog was used, in others a different individual of the same morph type. Although preparation of the second sample set followed the same protocol as the first (Etter et al., 2011), there were a few minor differences: First, to achieve higher DNA concentrations, I extracted samples using a standard ethanol-precipitation protocol, rather than a Qiagen kit. I confirmed the qualities of the extractions and the concentration of each sample using the Qubit system and a Nanodrop 2000, resulting in my exclusion of one F₁ offspring from family 6. Second, I sheared 1.0-1.85 µg of *Sbf*I-digested DNA from each pool using a Bioruptor sonicator for 10 30s intervals. Third, I used barcodes containing six, rather than five, nucleotides. Fourth, since the number of samples (85) exceeded the number of unique barcodes available (48), I split samples into pools such that barcode sequences were unique within pools but shared among them. This resulted in four pools: pool 1 (24 samples) barcoded with *Sbf*I 01-24, pool 2 (24 samples) barcoded with *Sbf*I 25-48, pool 3 (24 samples) barcoded with *Sbf*I 01-24, and pool 4 (13 samples) barcoded with *Sbf*I 25-37. To avoid demultiplexing errors, I ensured that barcodes differed by at least 2-bp. Fifth, I amplified pools using 16, rather than 18, cycles of amplification. I sequenced each amplification pool on one lane of an Illumina HiSeq2500, using 100-bp paired-end sequencing, at the FAS Center for Systems Biology at Harvard University.

1.3.5 Quality Filtering of Raw Reads

I de-multiplexed raw Illumina reads from both sequencing runs and filtered for quality using programs within the FASTX-Toolkit. I discarded low quality reads with fewer than 90% of bases assigned a Phred quality score > 10 . For reads containing five-digit barcodes (sample set one, described above), I removed reads with one or more mismatched barcode sequence. For reads containing six-digit barcodes (sample set two, described above), I excluded reads with >1 barcode mismatch. I combined into a single, filtered dataset the reads from wild-caught individuals that I included in both sample sets.

1.3.6 Population Genomics and Association Study

I used the STACKS pipeline version 1.19 (Catchen et al., 2011) to build loci, call SNPs within each locus, and generate within and between population and morph statistics. Using the STACKS program *ustacks*, I assembled the Illumina read 1 sequences from each wild-caught frog into unique stacks of identical RAD sequences, discarding any stacks with a read depth less than three. Next, I assembled the stacks into loci, allowing no more than three nucleotide mismatches between stacks in the same locus. To circumvent the issue of over-represented sequences, likely the result of repetitive elements (Rogers et al., 2018), I discarded any stack with a depth of coverage two standard deviations over the mean depth of coverage, and any stacks that differed from these deep stacks by just one nucleotide. I detected SNPs at each locus using a maximum likelihood framework. I assembled a catalog of consensus loci from all wild-caught individuals using the *cstacks* program, and matched individuals against this catalog with *sstacks*. Finally, I calculated within and among population and morph summary statistics (e.g., F_{IT} , F_{ST} , and

heterozygosity) using the *populations* program, excluding loci that were absent in more than 50% of individuals. I performed a Fisher's Exact Test to determine the significance of pairwise F_{ST} values between morphs of each population. I designated candidate coloration loci as any locus containing SNPs with a between morph $F_{ST} > 0.4$ and a p-value from the Fisher's Exact Test < 0.05 . I used a custom script (W. Zhang) that compared candidate coloration loci from both polymorphic populations to identify those shared between populations and red morphs.

Genotyping RAD Loci for Mapping

As with the association study of wild-caught frogs, I employed STACKS version 1.19 (Catchen et al., 2011) to the mapping study of laboratory colony frogs. I assigned genotypes to parents and offspring using a pipeline modified from that described in Catchen et al. (2011), because this dataset included loci from eight small families rather than one large family. Rather than process each family separately, I constructed the catalog of parental haplotypes from all 16 parents. This facilitated the identification of loci common to all families. I combined progeny loci into a 'superfamily' dataset and assigned genotypes by matching each locus with the parental catalog. In this way, I increased the maximum number of matching progeny for any given parental locus from 3-10 (the number of progeny in a family) to 50 (total progeny in the study). I compiled informative, mappable markers (parental loci heterozygous in one or both parents of a pair; Appendix Table 1), then exported them from STACKS for use in JoinMap.

Since not all markers were informative in the same families, and because some markers were genotyped in just a few offspring within a family, I used two of the largest families (families 1 and 2, Table 1) as indices to integrate markers across families. In this way, I excluded any markers that were found in smaller families but were not found in family 1 or family 2, resulting in 'reference 1' and 'reference 2' datasets. Next, I used the catalog of *O. pumilio* transposable

elements (Rogers et al., 2018) to isolate and then remove any RAD loci that mapped to these regions. I selected 500 markers with the least missing data for each segregation type. However, all markers were retained for mapping when fewer than 500 were available for that segregation type. I did this for both ‘reference 1’ and ‘reference 2’ datasets, resulting in a RAD marker dataset that included 1000 lmxll markers, 1000 nnxnp markers, 348 efxeg markers, 219 hkxhk markers, and 75 abxcd markers.

1.3.7 Amplifying Candidate Loci Regions in Pedigree Frogs

To amplify the candidate coloration loci in the pedigree frogs, I employed a primer design method that first used BLAST (version 35x1) to search the 88-bp RAD markers containing color-associated SNPs against an assembly of the *O. pumilio* genome (Rogers et al., 2018). This search isolated five RAD markers, corresponding to CL01, CL04, CL12, CL16, and CL17, that matched to unique scaffolds with 88-bp (100%) similarity. Two additional loci, CL10 & CL15, matched to unique scaffolds with 1 (99% similarity) or 2 (98% similarity) mismatches, respectively. The remaining RAD markers returned multiple hits, so I excluded them from further analysis. Next, I imported the RAD markers and their associated scaffolds into Geneious version 8.17, and designed primer pairs using Primer3 version 2.3.4. I validated each potential primer pair using the *primersearch* tool from the Emboss package, allowing no more than two mismatches between the primer sequences and the genome. I tested with PCR only the primer pairs that mapped to unique regions of the genome. Initial PCR using these primer pairs resulted in the amplification of genomic regions around five of the 18 color-associated RADs. Although my aim was to amplify the region including a color-associated RAD marker, there were no compatible primer pair

sequences on either side of the c10 marker. However, the amplified product sat less than 100-bp upstream from the RAD marker, so is still expected to contain a signal of color association.

I extracted fresh genomic DNA from toe-clips or tadpole tails in spin-columns using the Qiagen DNeasy Blood and Tissue Kit. I included the original 16 parents and their 49 F₁ offspring in these extractions, along with an additional 65 F₁ offspring from the same families. Each PCR reaction (12μL total volume) contained 1-2μL of genomic DNA template, 6μL of GoTaq Green Master Mix (Promega, Madison, WI), and 0.75-1.0μL of 10μM forward and reverse primers. I performed PCR using a SimpliAmp Thermal Cycler (Applied Biosystems, Foster City, PA). Thermal cycling conditions included an initial denaturation at 95°C for 90s, 35 cycles of 95°C for 30s, annealing at T_A for 45s, and an extension at 72°C for 45s, followed by a final extension at 72°C for 5 minutes. Primer sequences, annealing temperatures, and approximate amplified fragment length are given in Appendix Table 2. I confirmed amplification using gel electrophoresis, in which I ran 2μL of PCR product adjacent to a 100-bp DNA ladder (GenScript, Piscataway, NJ) at 100V, until bands were fully resolved. Following electrophoresis, I purified PCR products using ExoSap-IT (Affymetrix, Santa Clara, CA) and submitted for Sanger sequencing to GeneWiz.

1.3.8 Mapping Candidate Loci

I aligned parent and offspring sequences from each candidate locus region in Geneious, then genotyped informative markers for all individuals. I manually integrated this candidate region marker dataset, containing genotype scores for 16 parents and 114 offspring, with the RAD linkage mapping dataset (Appendix Table 3). I conducted a linkage analysis of the combined (RAD markers + candidate region markers) datasets in JoinMap4.1, calculating marker groups using the

independence LOD grouping parameter, beginning at LOD 1.0 and increasing incrementally by 0.1 to LOD 10.0. Finally, I generated linkage maps for selected marker groups using the maximum likelihood mapping algorithm with default settings.

1.4 Results

1.4.1 Population Genomic Scan

After filtering low-quality reads from both sequencing runs there was an average of 10.6 million Illumina read 1 sequences per wild-caught individual, corresponding to a mean of 148,771 loci per individual. 650 loci were shared among all 40 individuals. Eighteen loci passed all executed filters ($F_{ST} > 0.4$, $P < 0.05$, depth > 3 , coverage $> 50\%$) and are referred to as candidate coloration loci. Each of these 18 loci were manually validated in STACKS. Ten of these candidate coloration loci were found in red morphs from both Cemetery Hill and Dolphin Bay Preserve populations. Of these 10 ‘red loci’, four passed a strict filter ($F_{ST} > 0.5$, $P < 0.05$). Scatterplots of p values plotted against F_{ST} values for Dolphin Bay Preserve (top) and Cemetery Hill, Isla Bastimentos (bottom) are shown in Figure 2, illustrating the process by which candidate color loci were selected. BLAST returned proteins for three candidate loci (c01, c13, c14), although none are linked explicitly to color production. Table 2 lists the candidate locus ID, number of SNPs and offspring supporting each locus, and any associated BLAST results.

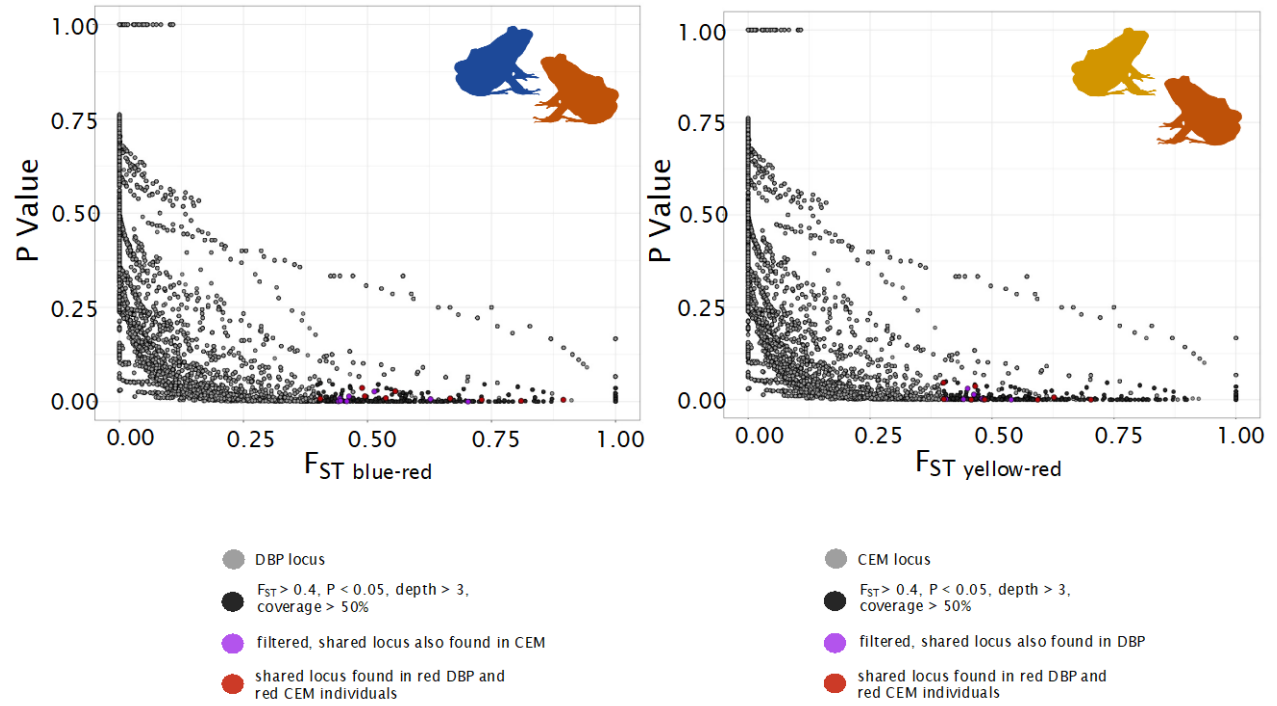


Figure 2. Scatterplots of p values plotted against F_{ST} values for Dolphin Bay Preserve (top) and Cemetery Hill, Isla Bastimentos (bottom).

Light gray datapoints indicate unfiltered loci and black datapoints indicate filtered loci. Purple datapoints indicate filtered loci that are found in both populations; red datapoints show filtered, shared loci found in red frogs from both populations.

Table 2. Candidate locus ID, number of SNPs and offspring supporting each locus, and associated BLAST results for 18 candidate coloration loci.

Candidate Locus ID	Number of SNPs	Number of Progeny	BLAST Results
c01	2	25	retinoblastoma-binding protein
c02	2	23	
c03	2	24	
c04	4	31	
c05	2	22	
c06	5	20	
c07	3	24	
c08	5	26	
c09	3	28	
c10	2	22	
c11	5	20	
c12	3	28	
c13	3	21	movement protein transposase
c14	1	20	
c15	4	23	
c16	4	26	
c17	2	28	
c18	4	24	

1.4.2 Linkage Mapping and Mapping Candidate Coloration Loci

After filtering low-quality reads there was an average of 3.7 million reads per laboratory frog, corresponding to 242,590 loci per individual. The initial attempt to generate a linkage map for this species estimated linkage groups from markers of all segregation types together, including maternally and paternally informative markers. Given the random digestion of genomic DNA during enzyme-based library preparations, and assuming even coverage across the genome, the distribution of markers across linkage groups should correlate with variation in chromosome size. Therefore, a starting point for the selection of linkage groups was the identification of the LOD threshold at which the number of markers across groups roughly corresponded with the karyotype for *O. pumilio* (Rassoto et al., 1987); LOD 2.3 for the combined dataset). Twelve of the 18 candidate loci markers were included in the first two groups at this threshold; the remaining six

candidates were in groups at lower thresholds that did not match the expected distribution. The first group (LG01, contains markers from all five of the candidate locus regions (c10_E, c12_E, c15_H, c04_O, and c16_E). The second group contains IM63 and IM58 from candidate locus c10_E, positioned at 140 and 142 cM, respectively. The occurrence of markers from the same candidate locus in different linkage groups (e.g., c10_E markers in LG01 and LG02), as well as the position of markers from the same candidate locus within a linkage group (e.g., in LG01 markers IM33 and IM22 from c15_H are separated from each other by RAD markers and three markers from c04_O), suggests that limitations of our experiment are precluding the inference of a sex-averaged linkage map.

Because of this, sex-specific maps were constructed using nearly identical settings in JoinMap 4.1. Fully (abxcd and efxeg) and partially (hkxhk) informative markers were recoded to lmxll and nnxnp genotype scores, creating male (lmxll) and female (nnxnp) datasets. Tentative linkage groups were tested at LOD thresholds 1.0 to 10.0 at 0.2 increments, resulting in groupings at LOD 2.4 for both male and female maps (Figs. 3,4). In the male map (Fig. 3), LG01 contains 169 markers, including candidate c16E_IM57; LG02 contains 123 markers, including seven candidate loci from four candidate regions (c04O, c10E, c12E, and c15H). The order of markers in these groups appear to be improved relative to the sex-averaged map, such that markers known *a priori* to be linked are grouped together. Map lengths for LG01 and LG02 are 1032 cM and 876 cM, respectively. In the female map (Fig. 4) LG01 contains 93 markers, including two candidate markers from c10E; LG02 contains 23 markers but no candidate loci. Map lengths for LG01 and LG02 are 611 cM and 64 cM, respectively. Although the largest group of 576 markers at an LOD threshold of 2.0 contained markers from c04O, c15H, and c10E, and seemed roughly equivalent to LG02 in the male map, the distribution of markers at this threshold did not match the karyotype

for *O. pumilio*. Furthermore, mapping of this large group (not shown) resulted in a very long map (5109 cM) in which the order of candidate markers did not match LG02 from the male map.

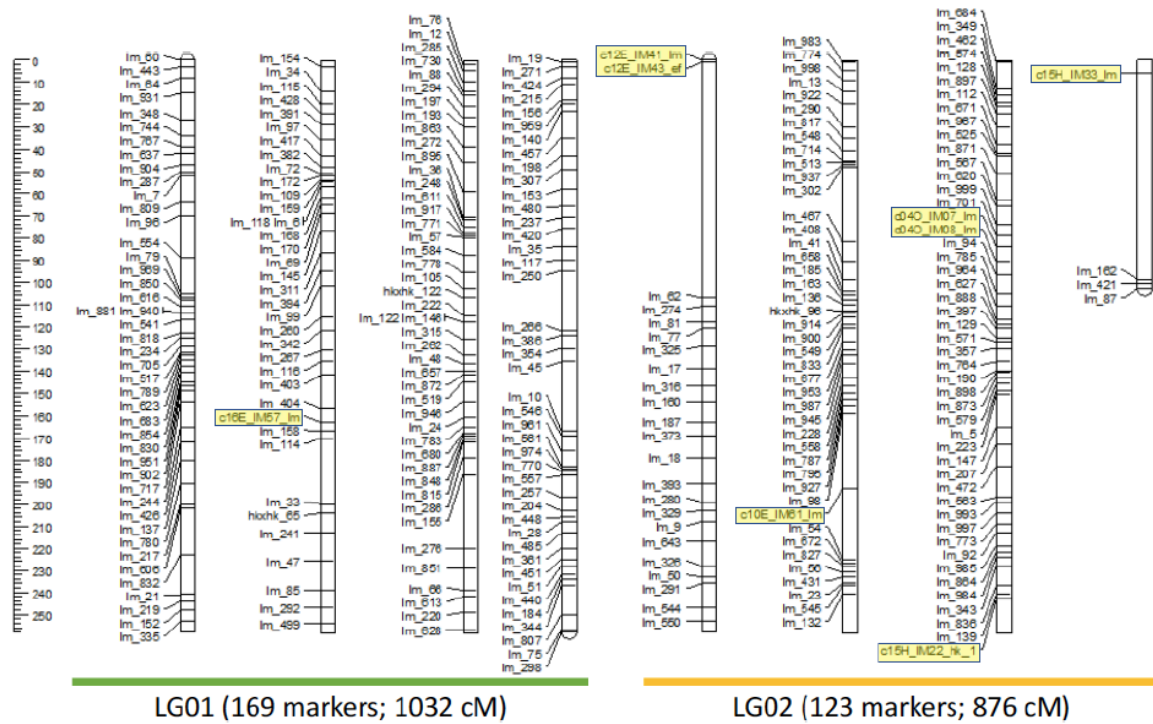


Figure 3. Sex-specific linkage map for male *O. pumilio* at LOD=2.4.

The largest linkage group, LG01, contains one candidate locus. The second largest group, LG02, includes seven candidate loci from four candidate regions (highlighted in yellow).

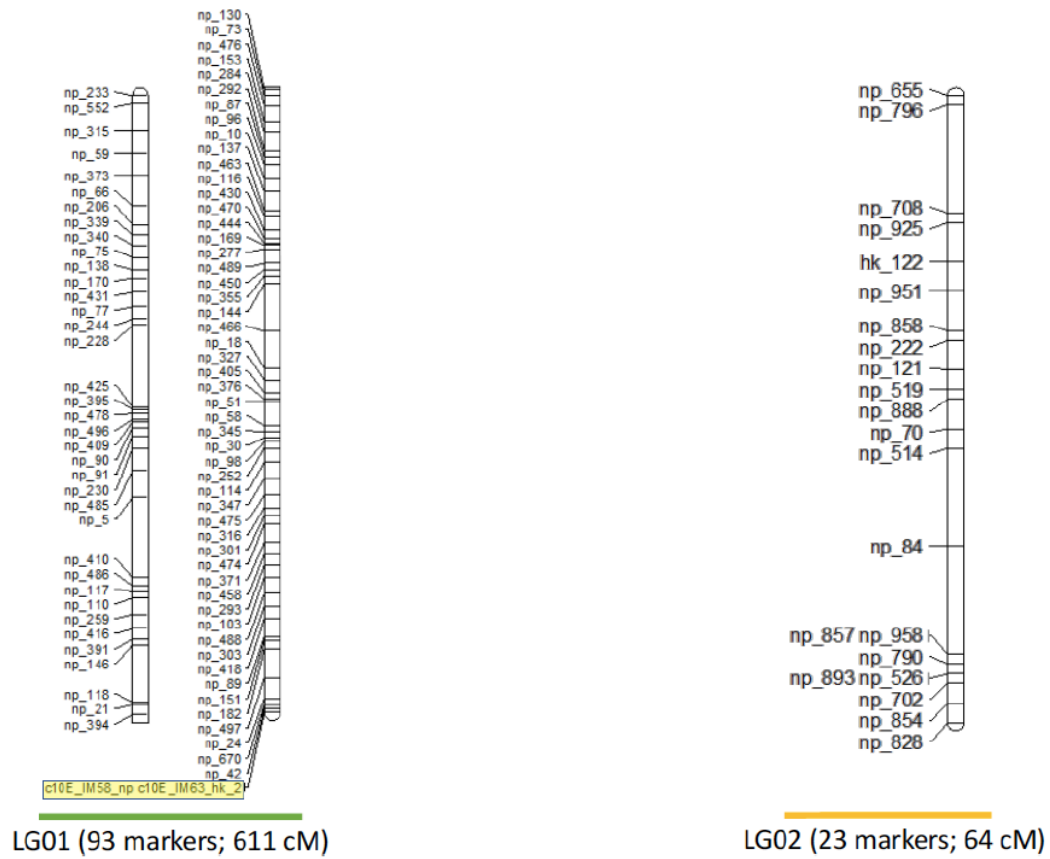


Figure 4. Sex-specific linkage groups for female *O. pumilio* at LOD=2.4.

The largest linkage group, LG01, contains two candidate loci from one candidate region (highlighted in yellow). The second largest linkage group, LG02, does not contain any candidate loci.

1.5 Discussion

Aposematic, or warning, coloration is a focal component of many well-studied animal systems and provides some of the most intriguing examples of adaptive evolution. Though some aposematic species present different geographical forms throughout their range (Briolat et al., 2019), few aposematic species display the remarkable level of phenotypic variation exhibited by

O. pumilio across the Bocas del Toro archipelago. Dramatic variation in dorsal background coloration, separation of morphs on islands, and a geologically young landscape make this species a fascinating system for studying the evolutionary processes leading to intraspecific divergence in aposematic coloration. A complete understanding of the factors leading to phenotypic divergence, in general, requires information from multiple levels of biological organization, from the genetic mechanisms underlying traits of interest to their ecological function (Brelsford et al., 2017).

Despite the application of next-generation sequencing and genomics to a diversity of non-model animals, only a handful (Stapley et al., 2010) approach a thoroughly integrated picture of evolution. These organisms are united by having small genomes, short generation times, and easily measured (and often discrete) phenotypic variation; a well assembled reference genome and ready captive breeding often accompany these features. However, a multi-level understanding of evolution should be the goal of every system, and biologists should not limit genomic applications to readily studied genomes. Species in which the function of phenotypic traits are known but currently lack an understanding of their underlying genetic mechanisms are especially good targets for genomic studies. The functional role of dorsal background coloration in *O. pumilio* has been examined in various evolutionary and ecological contexts (Gehara et al., 2013), positioning this species in a prime position for exploring the molecular basis of this phenotype.

With this in mind, my primary aim was to take a first but necessary step toward understanding the genetic basis of phenotypic divergence in dorsal background coloration among *O. pumilio* morphs. First, genomic variation was surveyed in wild-caught frogs from polymorphic populations and tested for significant associations between SNPs and dorsal background coloration. Next, five candidate coloration regions were genotyped in parents and offspring with known pedigrees, using Sanger sequences including and surrounding the candidate coloration loci.

Then sex-specific linkage maps were generated for *O. pumilio* using mappable RAD markers derived from the same pedigreed individuals. Candidate coloration regions were incorporated with this dataset to determine whether this extraordinary example of phenotypic variation implicates one or multiple genomic regions.

Using a combination of population genomics and traditional genetic mapping techniques, I identified multiple genomic regions that contain SNPs significantly associated with dorsal background coloration in *O. pumilio*. In the P1 (male) map, one of the five candidate locus regions map to a linkage group that may correspond with the largest *O. pumilio* chromosome (chromosome 1, Rasotto et al., 1987); the other four candidate regions span the extent of the linkage group that may correspond to the second largest chromosome (chromosome 2). In the P2 (female) map, the only mapped candidate locus resides on the linkage group that may correspond with chromosome 1. Together these results suggest polygenic control of dorsal background coloration in *O. pumilio*, though at this point it is not possible to determine whether it is due to few genes of large effect or many genes of small effect.

The capacity to investigate the molecular underpinnings of coloration in *O. pumilio* was somewhat limited by features of the RAD-seq method, sampling strategy, and species biology. RAD-seq and other reduced-representation approaches are valuable to identify genomic regions associated with traits of interest, particularly in organisms and phenotypes for which reference genomes or candidate genes are not available. It should be noted, however, that these approaches do not identify the genes that contribute to a phenotype. Instead, they identify genomic regions that are proximate enough to causal genes that they exhibit signals of linkage. It is possible, then, that I missed some of the loci associated with dorsal background coloration, particularly those located near genes of small effect.

The inclusion of additional morphs from other populations across the species range might reveal additional genomic regions involved in coloration or could even point to a different underlying architecture. Furthermore, I ignored color-associated loci that were not shared by both polymorphic populations to rule out spurious associations. Some of these excluded loci may be near color genes that contribute to differences between yellow and red morphs (in the Cemetery Hill population) or blue and red morphs (in the Dolphin Bay population). Follow-up studies with large sample sizes, coupled with data from morphological analyses, are required to explore loci unique to each population.

Perhaps the most severe limitations were imposed by *O. pumilio* life history and biology. A large genome (~6.3 GB; Rogers et al., 2018) results in lower coverage and fewer shared markers per unit of sequencing effort. Long generation times (~1 year) and difficult captive breeding limited the size of families in the linkage mapping to 3-10 individuals per parental pair. Smaller family sizes mean fewer recombination events and subsequently fewer mappable markers. These limitations manifest in the results as lower density linkage maps and ultimately the inability to generate sex-averaged maps.

Nevertheless, the findings here add to our understanding of the origin of diversified dorsal background coloration in *O. pumilio*. Demographic history will likely influence the level of selection required for phenotypic diversification. In the case of the strawberry dart frog, rapid population expansion into new environments meant that reduced levels of selection were required to undergo divergence in phenotype (Gehara et al., 2013). The genetic architecture of a trait can likewise influence the level of selection required for phenotypic diversification. The widespread distribution of candidate coloration regions means that stronger selection would be required than if these regions were linked, or even closer together on the same chromosome. So, although the

results of the present study suggest that strong selection was required to bring novel phenotypes to fixation, this barrier to divergence could have been lowered, or overcome, by the rapid expansion into new environments.

The results presented here open the door for further genetic research in this species. First, targeted sequencing and fine-scale mapping of the color-associated genomic regions can help identify candidate coloration genes or allow measures of the role of strong, recent selection on dorsal background coloration. Additionally, the SNP-based linkage map generated here- the first genetic linkage map of *O. pumilio* or for any dendrobatid frog- can complement association studies of other phenotypic traits of interest. Lastly, my results provide insight into the mode of sex-determination in *O. pumilio*. In amphibians, unlike mammals and birds, there are two types of genetic sex determination (XX/XY and ZZ/ZW). However, sex determination modes are classified for few species of amphibians (Duellman & Trueb, 1986). The shorter map lengths observed for females points to a ZZ/ZW system in *O. pumilio*, which is the presumed ancestral state for amphibians (Nakamura, 2009).

O. pumilio is just one of many frog species that utilize carotenoid pigments to achieve bright, and often conspicuous, coloration. Differences in carotenoid-based coloration among individuals or populations are attributed to differences in the genetic control of carotenoid metabolism, transport, and deposition. The genetic control of carotenoid-based coloration in amphibians is poorly understood (but see Twomey et al., 2020b). Furthermore, carotenoid pathways are thought to be highly variable compared to melanin pathways (Twomey et al., 2020). Given how coloration in amphibians is the result of several different pigment cell types containing melanins, carotenoids, or light-scattering structures (Bagnara et al., 1968), it is unsurprising that

dorsal background coloration in *O. pumilio*, and perhaps other brightly colored frogs, is under polygenic control.

In conclusion, this analysis provides an exciting first step toward understanding the genetic basis of dorsal background coloration in the strawberry poison frog, *O. pumilio*. My results suggest that the genes involved in variation in aposematic dorsal background coloration are located across broad regions of the genome, rather than being co-localized in one or few, nearby regions. This underlying architecture has implications for the potential for selection to result in divergent phenotypes. More specifically, the observed distribution of color-associated loci suggests that strong selection would be required to overcome the effects of recombination. Thus, my results set the stage for future investigations on the genetic basis of coloration and, more broadly, the evolution of phenotypic divergence.

2.0 The Proximate Mechanisms of Skin Color Variation in a Highly Polymorphic Poison Frog, *Oophaga pumilio*

2.1 Summary

Understanding phenotypic evolution requires considering all aspects of a trait, from how it is perceived by conspecifics and heterospecifics, to the underlying proximate mechanisms. Coloration is a complex trait, dependent on the morphological and functional assimilation of multiple components within the integument. The DCU is the functional unit of coloration in amphibians, birds, reptiles, and fish, and consists of reflecting structures and pigments that ultimately control the wavelengths of light emitted from the integument surface. Numerous studies examine how changes in the DCU result in differences in coloration, but few explore the mechanism of interspecific variation on the scale of *O. pumilio*. To bridge this gap, I explore the proximate mechanisms of coloration in 20 *O. pumilio* morphs. Using reflectance spectroscopy, histological analysis, and HPLC & UV/VIS spectroscopy, I examine the presence and absence of pigment types, ratios of pigments, extra-DCU components, and xanthophylls' chemical composition. Taken together, the results suggest that large color variations in *O. pumilio* are explained by differences in pigment proportions and carotenoid suites. I detected significant differences in the dermis, epidermis, and collagen layer heights between some color groups, but these components distinguish individuals by color to a much lesser extent than pigment proportions. In all but one population sampled, all three pigment cells were present, suggesting that the loss of pigment cell types is not a mechanism by which color change in this species commonly occurs. My results indicate that shifts from red to green coloration are accomplished

by decreasing reflecting platelets and mutations that suppress cis lutein ester or carotene in the skin. Although this study represents the early stages of understanding the proximate mechanisms of coloration in *O. pumilio*, the results motivate interesting future research topics.

2.2 Introduction

Animal coloration and patterning are some of the most variable phenotypes observed in nature. The intrinsically captivating diversity of animal coloration, whether it be of fur, feathers, scales, or skins, has provided biologists a multitude of systems to study phenotypic evolution. Several characteristics of animal coloration make it particularly well-suited to evolutionary studies. First, integument coloration is objectively quantifiable with relatively accessible technologies such as cameras and reflectance spectrophotometers. Second, coloration functions in intraspecific and interspecific interactions, and a physiological capacity (e.g., UV protection, evaporative water loss). Therefore, several selective forces, possibly acting in opposition, may shape this phenotypic diversity. Lastly, coloration can be highly variable among species and even populations, allowing for macroevolutionary and microevolutionary studies of its evolution.

Integument coloration in amphibians, and other poikilotherms such as fish and reptiles, is imparted by cells known as chromatophores. Chromatophores contain either 'true' pigments (biochromes) or structural colors (schemochromes; Taylor 1969; Morrison, 1995; Fujii, 2000) and are defined by their contents. Xanthophores contain yellow pteridines or red to yellow carotenoid pigments; the term 'erythrophore' can refer to chromatophores containing predominately carotenoids. Iridophores hold purines, structural pigments, arranged in stacks referred to as reflecting platelets (Taylor, 1966; Taylor, 1968; Frost et al., 1984). These cells appear silver or

white (Haslam et al., 2014), depending on the wavelength of light they reflect. Melanophores, generally the largest cells, contain brown-black melanins in organelles called melanosomes.

The spatial arrangement of xanthophores, iridophores, and melanophores in the skin is called the dermal chromatophore unit (DCU). This term refers to the pigment cells' physical and functional association in the dermis (Bagnara et al., 1968). Initially, the DCU concept explained the phenomenon of rapid (physiological) color change in amphibians (Taylor & Bagnara, 1972), a process by which pigments and schemochromes respond physiologically or morphologically to hormonal stimuli (Bagnara & Hadley, 1968; Taylor, 1968). While the DCU remains conceptually useful to explain physiological color change within individuals, it is also valuable to explore interspecific, intraspecific, and body region color differences (Taylor & Bagnara, 1972; Grether et al., 2004).

The DCU, along with mucous and granular (poison glands), nerve fibers, and smooth muscles, resides in the stratum spongiosum layer. A second dermal layer, the stratum compactum, consists of tightly packed collagenous fibers and lies directly below the stratum spongiosum. Unpublished experiments show that the stratus compactum emits blue color in the absence of chromatophores (Bagnara et al., 2007), suggesting this component can contribute to skin coloration. The epidermis, the most superficial skin structure, is separated from the dermis by the basal lamella (Duellman & Trueb, 1986). These three major components- the epidermis, stratum compactum, and stratum spongiosum- are identified easily with light microscopy (Fig. 5).

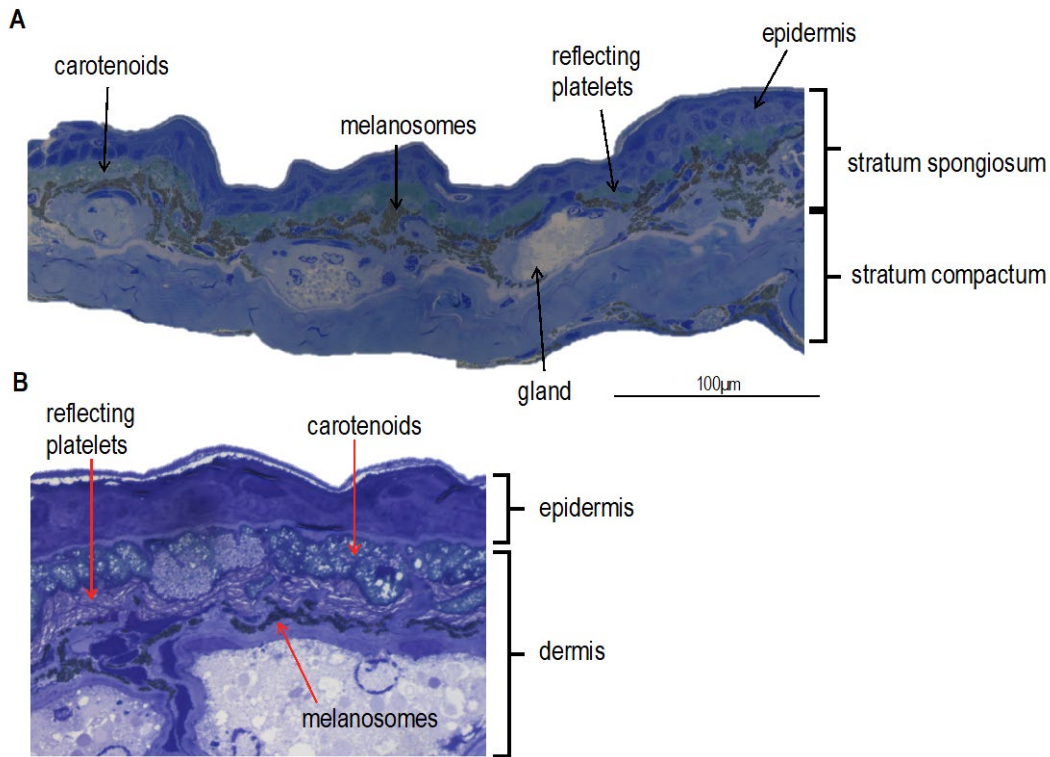


Figure 5. Stained histosections identifying major features of *O. pumilio* skin.

(A) The skin is organized into two main layers termed the stratum spongiosum and stratum compactum. The former contains the uppermost epidermis and underlying dermis. The latter consists of a compacted layer of collagen. (B) Micrograph showing the three-dimensional configuration of pigment cells in the dermis, termed the DCU throughout the text.

Observations of the vertical distribution of pigment cells in the DCU, and the wavelengths of light absorbed, reflected, or scattered by their contents, contributed to a general model of the basis of green skin coloration in amphibians (Bagnara & Hadley, 1973; Nielsen & Dyck, 1978). In this simplified model, xanthophores are the DCU's uppermost layer and absorb most short (blue-violet) wavelengths. Iridophores lie underneath xanthophores and reflect medium (yellow-green) wavelengths and any remaining short wavelengths. Melanophores, residing in the DCU's most basal location, absorb long (red-orange) wavelengths. Since the xanthophores re-absorb any blue-

violet wavelengths reflected by the iridophore layer, the viewer observes only the remaining yellow-green light.

The ultrastructural basis for the generation of bright (red, orange, yellow) or blue coloration is not understood as well, particularly in the context of a single, polymorphic species. However, previous studies in amphibians and other vertebrates demonstrate the numerous ways in which the DCU can differ among and within individuals, resulting in variation in the wavelength of light emitted (i.e., what the viewer perceives as skin color variation). One of the most substantial modifications is a complete absence of pigment cell types from the DCU. In the nocturnal leopard gecko, *Eublepharis macularius*, the skin contains xanthophores and melanophores, but no iridophores (Szydlowski et al. 2015). Likewise, male stony creek frogs (*Litoria wilcoxii*) brown and yellow dorsal skins, lacked iridophores in yellow and brown skins patches (Kindermann & Hero, 2016).

The ratio of pigment cell types can also vary. In the African frog, *Chiromantis petersi*, Drewes et al. (1977) observed iridophore layers as thick as 3-5 cells, many more than the typical one-layer thick, and proposed this as an adaptation to arid climates. In the tree frogs *Hyla cinerea* and *H. arenicolor*, there are as many as 10-15 iridophores for every melanophore (Berns & Narayan, 1970). Similarly, in the spotted yellow-spotted salamander, *Ambystoma maculatum*, each yellow xanthophore is accompanied by groups of iridophores. The outcome is a greater reflection of light from the reflecting platelets back through the xanthophore layer, resulting in increased yellow intensity (Taylor & Bagnara, 1972).

Other morphological differences may result in shifts in coloration. In iridophores, platelet size, spacing, and orientation determine how they interact with incoming light (Bagnara & Ferris, 1971; Taylor & Bagnara, 1972). Platelets oriented at different angles in the iridophore, as observed

in *H. cinera* and *H. arenicolor*, potentially scatter light (Taylor, 1969). In contrast, platelets with a uniform distribution, as in *A. dachnicolor* (Taylor, 1969; Taylor & Bagnara, 1972), most likely reflect incoming light.

Differences in the suite of yellow-red pigments contained within xanthophores provide a particularly fascinating mechanism for color variation. From a biochemical perspective, pterins and carotenoids are acquired from entirely different processes. In animals, pterins are metabolically derived from guanine triphosphate (Kim et al., 2006), whereas carotenoids are environmentally derived from their diet. Xanthophores may contain pteridines, carotenoids, or both stored in separate vesicles (Fujii, 1993). For carotenoid pigments, a series of complex biochemical processes are involved in the transport and deposition into the skin (Toews et al., 2017). Furthermore, many dietary carotenoids are modified and converted into ketolated forms, requiring additional biochemical pathways (Weaver et al., 2016). Since skin carotenoids are energetically expensive to produce, they are considered honest signals of mate condition (Endler, 1980).

Relative concentrations of pterins and carotenoids, and the diversity of each, influence the xanthophore's specific absorption range (Grether et al., 2005). Carotenoids are a particularly diverse class of organic molecules with hundreds of described pigments (Britton, 1995). Since changes in chemical structure will alter their absorption properties, carotenoids exemplify how small variations within the DCU result in shifts in overall skin coloration.

Finally, structures outside the DCU, such as the epidermis and stratus compactum, can alter total reflectance. At least one study reports differences in epidermis thickness between sexes and body regions (Greven et al., 1995). Although a thicker epidermis may confer physiological adaptations, it also increases the distance that light travels before reaching the DCU. Likewise,

there are morphological differences in the stratus compactum across body regions (de Birtó-Gitirana & Azevedo, 2005), potentially influencing the wavelengths scattered by this layer. Despite these observed variations in epidermis or stratus compactum morphology, no published studies examine their effect, if any, on amphibian skin coloration. Moreover, whether these differences are a mechanism for color variation or a byproduct of physiological adaptations is also untested.

Assuming evolutionary processes act only on the whole DCU, rather than individual structures or pigments, will unnecessarily restrict the hypotheses tested (Grether et al., 2004) and limit our understanding of the proximate mechanisms of skin color. To this end, Grether et al. (2004) argue for a multi-trait approach to the study of coloration. When we account for selection acting independently on DCU components, we have a mechanism by which the integument can respond to opposing evolutionary forces.

Poison dart frogs (family *Dendrobatidae*) have long garnered evolutionary biologists' attention, in no small part because of their eye-catching skin coloration. The strawberry poison frog, *Oophaga pumilio*, is a highly polymorphic Dendrobatid native to Central America (Batista & Köhler, 2008). In the region of Bocas del Toro, Panama, dorsal skin coloration varies from bright reds, oranges, and yellows to dull greens and blues (Roberts et al., 2006). Previous studies examining the selective forces at work in this system have implicated natural and sexual selection (reviewed in Gehara et al., 2013). The evolutionary processes that shape polymorphic coloration in this species are complex and underscore the need to understand their proximate mechanism.

In the present study, I explore the proximate mechanisms of interspecific variation among *O. pumilio* color morphs from Bocas del Toro, Panama. Using morphological measurements I collected from light micrographs and chemical data from high-performance liquid

chromatography (HPLC), I address the following questions: (1) Are color differences explained by the absence of pigment cell types? (2) Is color variation explained by differences in the ratio (proportion) of pigment cell types? (3) What role does variation in the combinations and suites of carotenoids or pterins play in *O. pumilio* explaining skin color? (4) To what extent does epidermis, dermis, and collagen thickness (height) affect skin coloration?

2.3 Methods

2.3.1 Sampling Strategy for Reflectance Spectroscopy and Morphological Analysis

Sampling for reflectance spectroscopy and morphological analysis included 20 populations representing all the color morphs described from Bocas del Toro archipelago and the adjacent mainland (Fig. 6, Appendix Table 4), plus one ancestral Costa Rican population. I collected five adult frogs per morph from each collecting site (n=5 monomorphic sites, n=15 Dolphin Bay Preserve, n=10 for Cemetery Hill, Bastimentos). I transported and temporarily housed individuals together in plastic containers until further processing at the Smithsonian Tropical Research Institute (STRI), following IACUC protocol No.15106566 (University of Pittsburgh). I photographed the dorsal and ventral bodies using a Panasonic DMC-TS5 camera affixed to a photo stand, with each frog positioned on a gray card, adjacent to a color standard and ruler.

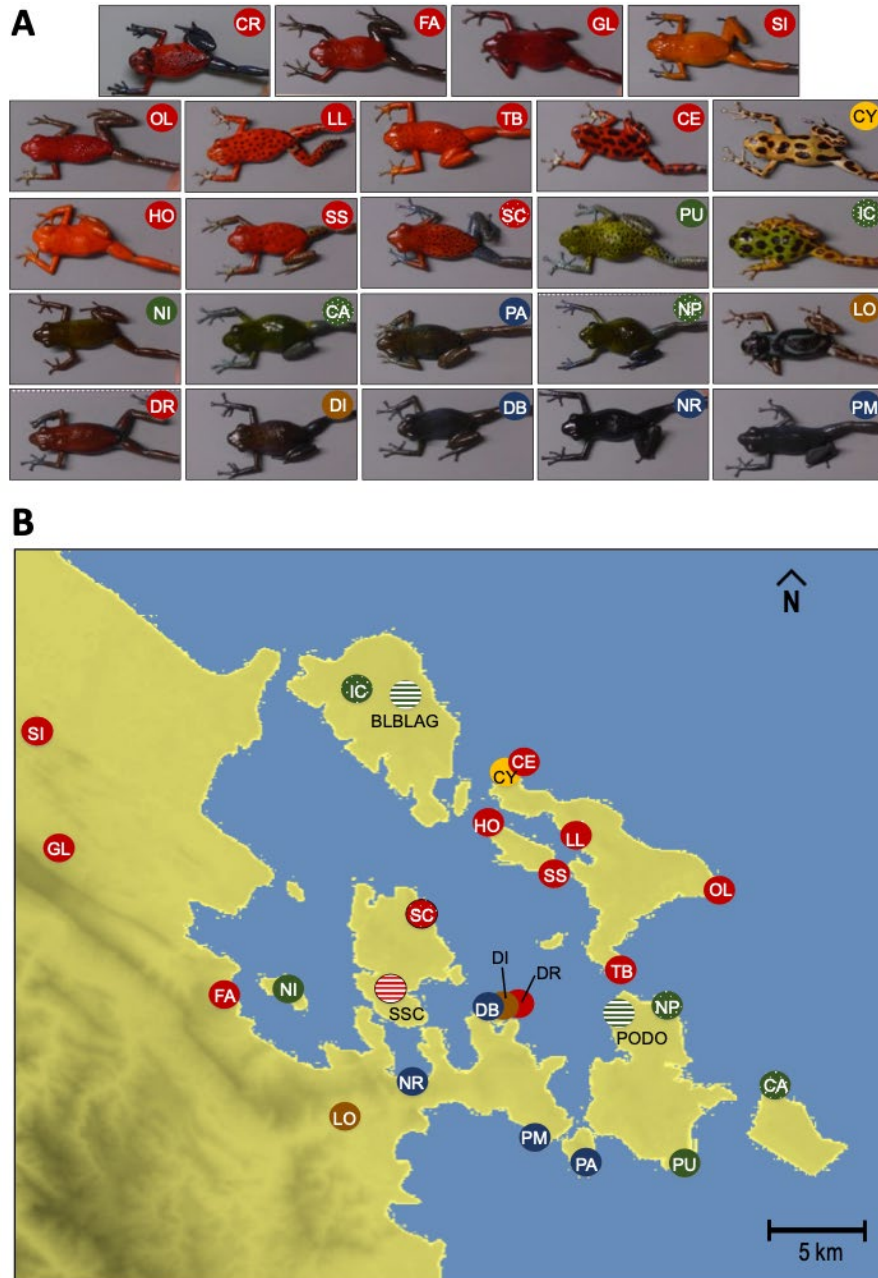


Figure 6. Sampling strategy for histological and chemical analyses.

(A) Pictures illustrating the diversity of color morphs included in this analysis. Colored labels in the upper right corner of each photograph correspond to sampling location labels in the map. (B) Map of Bocas del Toro archipelago in Panama. The positions of colored points represent sampling locations and are labeled with population/morph abbreviations. Points are color-coded according to by-eye dorsal skin coloration.; solid points indicate populations/morphs included in histological and chemical datasets, striped points indicate HPLC-specific samples, and dotted points indicate histological-specific samples. Polymorphic

populations are indicated by partially overlapping points. The ancestral Costa Rican morph is not shown here.

2.3.2 Reflectance Spectroscopy and Color Assignment

I measured dorsal and ventral coloration at three predefined regions per side with a JAZ-A1507 reflectance spectrophotometer connected to Oceanview spectroscopy software. I took measurements in a darkroom and calibrated the instrument at each session's start with light and dark standards. Real-time intensity-vs-wavelength plots ensured that I avoided data collection at melanistic spots. Upon completing one round of measurements for a set of five frogs, I randomized the sample order, recalibrated the spectrophotometer, and repeated the process a second time for each frog.

I processed raw spectral data, visualized spectra, and then calculated colorimetric variables using *pavo* in *R* (Maia et al., 2019; R Core Team, 2020). I restricted spectra to the visible range, 300-700nm, and averaged wavelengths into 1nm bins. To address a technical artifact that resulted in upward or downward 'blips' at three distinct 3-5nm ranges, I linearly interpolated the wavelengths between the artifact's start and end. I applied a minimal smoothing parameter of 0.05 to the spectral measurements integrated over less than 800ms (twenty individuals from 4 populations).

I used the *summary.rspec* function in *pavo* to calculate total brightness (B_1), the wavelength at the reflectance midpoint (hue, H_3), and the relative contribution of a spectral range to total brightness (chroma, S_1). These commonly used colorimetric variables are described in detail in Montgomerie (2006). *pavo* splits S_1 into six wavelength ranges corresponding with human/trichromat-perceived hues (Fig. 7): S_{1U} (λ min-400nm, excluded from this study);

S_{1V} (lambda min-415nm); S_{1B} (400nm-510nm); S_{1G} (510nm-605nm); S_{1Y} (550nm-625nm); S_{1R} (605-lambda max). These three variables- B_1 , H_3 , and S_1 - provide an objective means to quantify coloration.

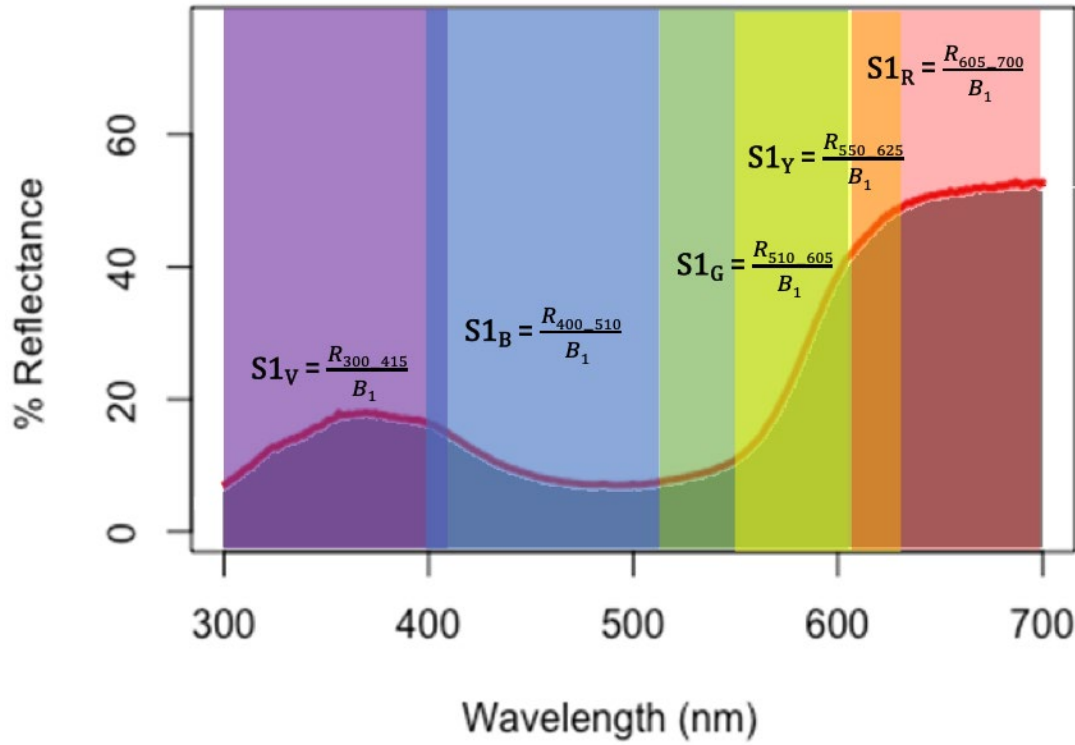


Figure 7. A reflectance plot illustrating the color metrics used in this analysis.

Total brightness (B_1) is a sum of the area under the red curve. Each chroma metric (S_{1V} , S_{1B} , S_{1G} , S_{1Y} , and S_{1R}) is calculated from the proportion of B_1 in its corresponding wavelength range. The formulas for each metric are shown for each segment.

To assess repeatability between measurement rounds one and two, I calculated R^2 values from scatter plots of B_1 at each predefined region (e.g., dorsal spot 1 round 1 vs. dorsal spot 1 round 2). To identify errant spectral measurements, I visually inspected spectra and validated outliers with corresponding frog photos, calculated individual-level variances in B_1 , and examined scatter plots of B_1 for pairs of regions (e.g., B_1 /dorsal spot 1/round 1 vs. B_1 /dorsal spot 2/round 1).

By compiling these techniques, I identified and then removed outlier measurements resulting from technical error rather than natural color variation.

I assigned individuals to color categories using model-based hierarchical cluster analysis (HCA) of S_1 (chroma) metrics that I averaged across individuals. This cluster analysis method, implemented in the *mclust* package in *R* (Scrucca et al., 2016), allows users to identify the most likely multivariate mixture model and number of clusters using maximum likelihood and Bayesian Information Criterion (BIC). First, I performed separate cluster analyses of dorsal and ventral data using the first two principal components of S_1 metrics. Next, I chose a covariance model based on BIC values and an examination of individual-level classification uncertainty. Finally, I assigned individuals, and then populations, to color groups. For the ventral coloration cluster analysis, I ultimately excluded data derived from the region closest to the vent (V03) as it often shares coloration with the legs and misrepresents ventral coloration.

2.3.3 Dorsal Skin Morphology

To explore how morphological differences in the dermal chromatophore unit can affect frog coloration, I quantified pigment areas, heights, and dermis areas from light micrographs of dorsal integuments. First, I euthanized frogs via decapitation, followed by double-pithing, according to IACUC protocol (University of Pittsburgh, No.15106566). I dissected the dorsal skins from the body and immediately placed them in separate glass vials containing ~1.5 mL of room temperature Karnovsky's fixative (Karnovksy, 1965) at pH 7.0-7.2. After ~2h, I replaced the fixative with 0.2M sodium cacodylate buffer (pH 7.2) and stored the vials at 4°C until transportation on dry ice to the University of Pittsburgh.

Using a stereomicroscope, I excised up to four 1-2mm rectangular pieces ("matchsticks") from each skin sample, avoiding any melanistic spotting areas. After postfixation with 1% osmium tetroxide in 0.2M sodium cacodylate buffer, I dehydrated matchsticks in an ethanol series followed by propylene oxide rinses, embedded each matchstick in epon, and then cured blocks at 60°C for a minimum of 48h. Histologists at the Center for Biological Imaging (CBI) at the University of Pittsburgh prepared thick (1 μ) unstained and thin (350nm) stained sections with an ultramicrotome and affixed sections to glass slides.

Using a Zeiss TissueGnostics brightfield scanning scope connected to a PixeLink PL-B622CF microscope camera and TissueFAXS Imaging Software, I scanned high-resolution images of one histosection per matchstick at 400x total magnification. I repeated image collection of the unstained thick sections, without repositioning the slides, using a custom-fitted polarizer. This procedure guaranteed that I could later overlay standard light images of the entire histosection with corresponding polarized light images of the reflecting platelets.

I utilized RGB and single-channel color thresholding, freeform, or straight-line measurement tools within NIS-Elements and ImageJ software (Schneider et al., 2012) to quantify the following morphological features:

1. Epidermis, dermis, collagen layer, and total section heights from three subregions
2. Total histosection length
3. Dermis area, minus glands
4. Areas of melanins, carotenoids, and reflecting platelets in the dermis

I then calculated from each histosection the proportions of total pigment (any pigment), melanosomes, carotenoids, and reflecting platelets in the dermis. The final morphological dataset contained measurements from 1-4 histosections per individual and 3-5 individuals per morph.

2.3.4 High-performance Liquid Chromatography (HPLC) and Ultraviolet-visible (UV/VIS) Spectroscopy of Carotenoid Pigments

Sampling for HPLC and UV/VIS spectroscopy included 16 populations from the Bocas del Toro archipelago and adjacent mainland (Fig. 6; Appendix Table 4). I collected five to 13 adult frogs per morph from each collecting site, then transported, housed, photographed, and euthanized frogs at STRI in the manner described above. I immediately removed dorsal and ventral skins, then flash-froze them in liquid nitrogen. The samples remained frozen during and after transportation to the United States, either in liquid nitrogen, at -80°C , or on dry ice, until further processing.

The HPLC and UV/VIS spectroscopy analysis required a minimum sample mass of 0.2g per sample, so I pooled individual frogs' skins by color morph and population ($N = 5$ to 13 individuals per pooled sample). However, in cases where the dorsum and venter were differently colored, I pooled individuals but kept dorsal and ventral skin pools separate for analysis. As color is not usually sexually dimorphic in this species (but see Mann & Cummings, 2009), both male and female frogs contributed to each pool.

The analyses were performed on pooled samples in Kevin McGraw's lab at Arizona State University. McGraw et al. (2006) detail the analytical methods for HPLC, and Crothers et al. (2016) describe modifications made for *O. pumilio*. In brief, the McGraw lab conducted a pterin assay of skin samples with a 1% NH_4OH solution and quantified the extract utilizing UV/VIS spectroscopy; carotenoids were extracted with a 1:1 mixture of Hexane/Methyl Tertiary Butyl Ether (MTBE) and quantified using a Waters HPLC 2695 with PDA detector 2996 (HPLC). Each sample pool returned a chromatogram with the absorbance wavelength (λ) over retention time (min). The values of peak absorbances (λ_{max}) in the chromatogram determines the identities of pigment types, and the area under each peak estimates the quantity of each pigment type (in $\mu\text{g/g}$).

2.3.5 Compositional Treatment of Proportional Pigment Data

I analyzed the proportional morphological data within a compositional data analysis (CoDa) framework. In CoDa, data are log-ratio transformed from a vector space on the simplex to Euclidean space, allowing for analysis by classical multivariate statistics (Boogaart & Delgado, 2013; Filzmoser et al., 2018). Interpreting results may require back-transformation, mainly if it requires the original proportional data. More recently, studies in diverse fields producing compositional datasets have adopted this general workflow, referred to as 'the principle of working in coordinates' (Mateu-Figueras et al., 2011).

I first applied the Aitchison (i.e., ratio) compositional scale using the package 'compositions' (van den Boogaart & Greifswald, 2005) for *R*. This scale is appropriate for datasets in which the individual amounts are part of a whole and closing the data is inconsequential (Aitchison, 1982). Next, I explored the presence of compositional outliers using the *OutlierClassifier1* function. The only outliers identified by the function were pigment proportions measured from (by-eye) blue frogs, so I ultimately left these in the analysis so as not to exclude real, biological variation.

I observed several histological sections in which reflecting platelets or carotenoids were completely absent. This finding is not unprecedented, so I treated the '0' observations for these pigments as informative ('essential' or 'true' zeros). As opposed to missing values, true zeros cannot be analyzed under the CoDa framework because log-ratio transformations fail with zero values. Therefore, I substituted zeros with minimal values and increased every non-zero observation by the same amount (pseudocounts; iridophores = 0.00001; carotenoids = 0.001).

Finally, I transformed the pseudocount-adjusted dataset with a specialized isometric log-ratio transformation known as pivot coordinates (Egozcue et al., 2003). In pivot coordinates of a

three-part compositional dataset, all the relative information about two compositional parts (e.g., melanosomes and carotenoids) is aggregated in the first coordinate (e.g., reflecting platelets). This strategy allows for separate statistical models of each component, thereby teasing apart its contribution relative to others.

2.3.6 Statistical Analysis of Morphological and Chemical Data

To examine the effects of pigment abundance on dorsal skin coloration, I modeled the relationship between the proportion of pigmented dermis and colorimetric variables (brightness or chroma) in a linear-mixed effects framework, as implemented in the *nmle* (Pinheiro et al., 2020). I applied logarithmic or arcsine transformations to non-normally distributed color variables (all but Sl_Y). In these models, I treated total pigment as the response variable, color metrics as the explanatory variable, and frog ID and histological section ID as nested random effects (i.e., total pigment \sim color variable, random= ~ 1 |frog ID/histological section ID). I calculated marginal and conditional R^2 values, the latter of which considers the variance by both the fixed and random effects, using the *rsquared* function in *piecewiseSEM* (Lefcheck 2016).

I employed the same linear-mixed effects framework to investigate the relationship between pigment proportions and dorsal colorimetric variables (brightness or chroma). I retained individual- and histosection-level pigmentation data but averaged colorimetrics across individuals to simplify the model structure. As in the pigment abundance models described above, I used transformed explanatory variables, then ran separate models for each compositionally-transformed pigment component. In each model, pigment component is the response variable, the colorimetric variable is the explanatory variable, and frog ID and histological section ID are nested random effects (i.e., pigment component \sim color variable, random= ~ 1 |frog ID/histological section ID).

This design produces three models for each color variable, one for each pigment type (reflecting platelets, melanosomes, and carotenoids). As with the pigment abundance models, I obtained R^2 values with *piecewiseSEM*.

To test if dermis, epidermis, or collagen heights are different among color groups, I performed a multivariate repeated measures analysis of variance (MANOVA). For epidermis and collagen heights, I averaged measurements from three subregions per histosection. However, I obtained mean dermis height by dividing the dermis area by section length, since this included the entire histosection. After testing model assumptions and removing multivariate outliers, I ran the analysis with mean heights from 1-3 histosections per individual as dependent variables and the *mclust*-derived color assignments and histological section number as independent variables.

I compared linear discriminant function models with and without height data to investigate how well pigment proportions or epidermis, dermis, and collagen layer heights separate individuals by color group. After testing model assumptions and removing outliers, I performed robust linear discriminant analysis (LDA) in the *MASS* package for *R* (Venables & Ripley, 2002). In these models, I used pivot coordinates for the pigment proportions and log-transformed height data to find the discriminant function that maximizes separation among the four *mclust*-derived color groups. Prior probabilities were derived from the observed frequencies of histosections in each color group, and I used standard estimators of mean and variance. I applied a resubstitution to predict color group membership, then compared prediction accuracy of pigment proportions with and without heights data. To account for accuracy inflation with this method, I re-estimated prediction accuracy with leave-one-out cross-validation approaches. I visualized the discriminant functions and classification results with the *partimat* function from the *klaR* package (Weihs et al., 2005).

To explore if the composition of carotenoids are different among color groups, I applied non-metric multidimensional scaling (NMDS), analysis of similarities (ANOSIM), and permutational multivariate analysis of variance (PERMANOVA) to HPLC abundance data from dorsal skins. I applied posthoc similarity percentages breakdown (*SIMPER*) tests (Clarke, 1993) to identify which carotenoids drive differences between pairs of color groups. Since I derived color groups using the first two principal components of the five S_1 metrics, I performed Mantel tests of total brightness (B_1), the S_1 metrics combined, and S_R , S_Y , S_G , S_B , and S_V . This procedure allowed me to compare the correlation between carotenoid composition and individual colorimetrics. Based on the *SIMPER* and *MANTEL* results, I examined the strength and direction of the relationship between single carotenoid pigments and colorimetric variables (e.g. apocarotenoid and S_{1G} , respectively) using linear models.

Finally, to assess the agreement between my morphological and HPLC datasets, I implemented a linear-mixed effects framework to model the relationship between the log-transformed total carotenoids from HPLC and (1) the log-ratio transformed proportion of carotenoids measured from histosections or (2) B_1 . If my datasets agree, I expect to see strong, positive relationships for both models, such that total carotenoids measured from HPLC increase with B_1 and the proportion of carotenoids measured from histosections.

2.4 Results

2.4.1 Characterizing Coloration from Spectral Reflectance Data

To provide objective measures of coloration, I collected and analyzed reflectance data from the dorsum and venter of each frog. Figure 8 represent the results of model-based cluster analyses of chroma metrics for dorsal and ventral data. For the dorsal data, the best model separates data into four components, with a within-group covariance matrix assuming an ellipsoidal distribution, equal shape and volume, and variable orientation (EEV; BIC = 337.7; Fig. 8A). Three of the four clusters correspond closely with by-eye color assignments: cluster 1 (green frogs), cluster 2 (red frogs), and cluster 4 (blue frogs). A fourth cluster (cluster 3) contains brown/'other' frogs from several blue and green morphs. I visually inspected cluster plots for other highly supported models, the EEV density plot with four multivariate normal distributions, and an uncertainty plot for the EEV model, and decided this cluster model was most appropriate for my data, despite one of the clusters disagreeing with by-eye assignments.

For the ventral data, the best model separated data into five components, with a within-group covariance matrix assuming an ellipsoidal distribution, variable shape, equal volume, and equal orientation (EVE; BIC = 440.6; Fig. 8B). However, two clusters corresponded to red frogs, with two populations/morphs present in both. For this reason, I chose the third-best supported model, which separates ventral data into four components. This model assumes an ellipsoidal distribution, variable volume and shape, and equal orientation (VVE; BIC=438.0). All four clusters corresponded closely with ventral by-eye color assignments: cluster 1 (green), cluster 2 (white), and cluster 3 (red), and cluster 4 (blue).

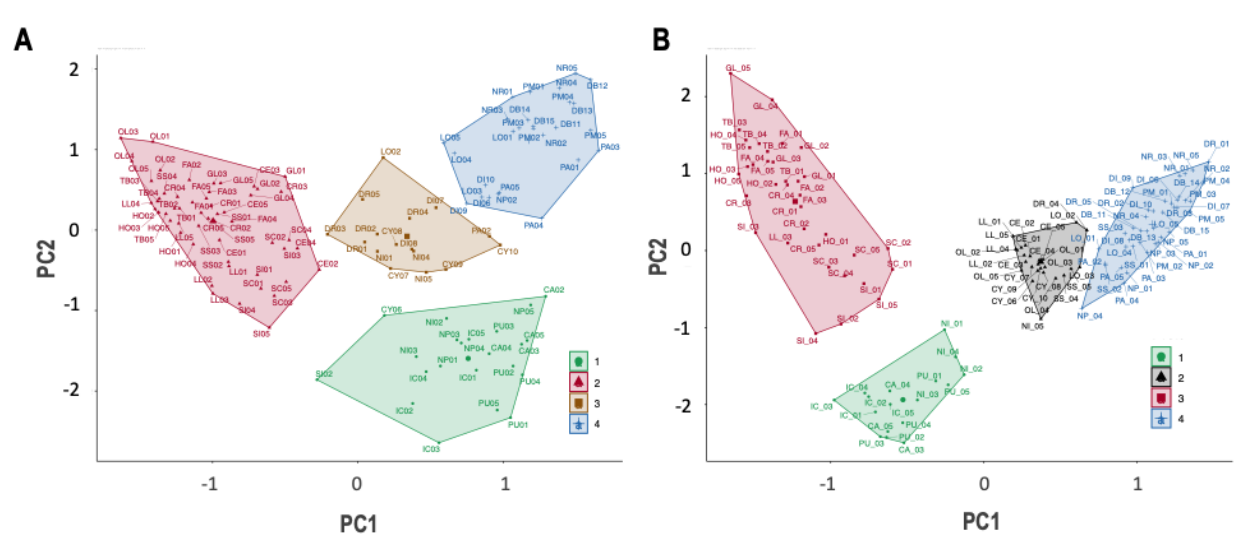


Figure 8. Visualization of dorsal (A) and ventral (B) clusters based on model-based cluster analyses.

Data points are labelled with population/morph IDs, which correspond to abbreviations used in Figure 6, and individual frog IDs. Clusters are colored according to by-eye color assignments, except for dorsal cluster 3 (brown/'other') and ventral cluster 2 (white venters).

2.4.2 Morphological and Chemical Determinants of Skin Coloration

2.4.2.1 The absence of pigment types from individuals and morphs

Previous studies in amphibians and other vertebrates report the absence of pigment cell types from the dermis of some species or morphs. To this end, I examined multiple histosections from most individuals, and multiple individuals per morph. Every histosection I examined contained dermal melanosomes. However, two histosections from two morphs lacked reflecting platelets, and numerous samples showed no evidence of carotenoids. I did not find carotenoids in any histosections from one blue morph (Loma Estrella, 8 histosections from 3 individuals). In the other morphs for which I observed histosections without carotenoids (blue frogs from Rana Azul, La Loma Partida, and Partida Mainland), I found small amounts of carotenoids in other

histosections or individuals. These results suggest that the absence of a pigment type in *O. pumilio* morphs is unique to the Loma Estrella (LO) blue morph.

2.4.2.2 Total pigment and pigment proportion effects on chroma, total brightness

To investigate if the total proportion of pigmented dermis or individual pigment proportions predicts colorimetric variables, I modeled their relationships in a linear-mixed effects framework. Based on these models, total pigment does not predict any colorimetric variables (Appendix Table 5). However, reflecting platelet, melanin, and carotenoid proportions predict most colorimetric variables (Table 3, Figs. 9, 10).

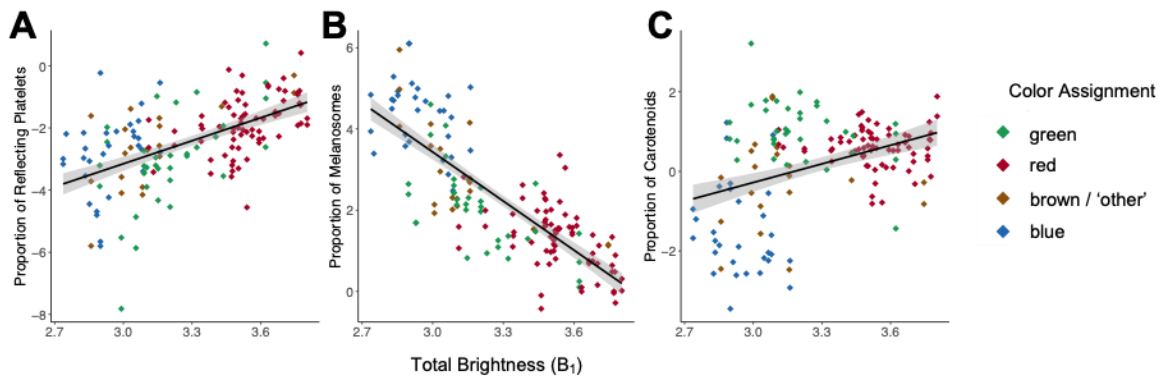


Figure 9. Scatterplot of pivot coordinates of pigment proportions versus total brightness (B_1).

Regression lines indicate significant relationships among response and explanatory variables.

There are positive relationships between reflecting platelet and carotenoid proportions with total brightness, and a negative relationship between melanin proportion and total brightness (Fig. 9). These results are unsurprising in light of previous studies in other vertebrates.

The relationships between pigment proportions and chroma metrics are more complex than total brightness. Taken together, the linear mixed-effects models imply the roles of reflecting platelets, melanins, and carotenoids shift from as you move along the visible spectrum from short

to long wavelengths (Fig. 10). In the blue and violet (short) wavelength ranges, chroma increases are associated with an increase in melanins, but decreases in reflecting platelets and carotenoids. On the other end of the visible light spectrum (oranges and reds, long wavelengths), chroma increases are associated with a decrease in melanins, but increases in reflecting platelet and carotenoid proportions. Changes in chroma in the medium wavelength range (510-605nm; green) are correlated only the proportion of reflecting platelets, suggesting other morphological features determine green coloration.

Table 3. Summary of results of linear mixed effects models with pivot coordinates of pigment proportions as the predictor variables and color metrics as the explanatory variables.

Explanatory Variable (x)	Response Variable (y)	Fixed Effects Slope Value	Standard Error	p-value	AIC	Marginal R ²	Conditional R ²
B_I	<i>ref. platelets</i>	2.37013	0.3461413	< 0.0001	475.5266	0.2869725	0.9407468
	<i>melanins</i>	-3.98531	0.3102211	< 0.0001	375.6028	0.6341492	0.9769501
	<i>carotenoids</i>	1.698644	0.3924027	0e+00	441.8932	0.1660658	0.9495027
S1_v	<i>ref. platelets</i>	-5.953662	0.9909820	0.0000	480.4658	0.2358752	0.9369682
	<i>melanins</i>	10.985270	0.8246706	0e+00	369.1716	0.6389196	0.9768174
	<i>carotenoids</i>	-5.307180	1.0311161	< 0.0001	433.6	0.2103214	0.950753
S1_B	<i>ref. platelets</i>	-1.045541	0.2778684	3e-04	499.1232	0.1135351	0.9303518
	<i>melanins</i>	2.723003	0.2150842	< 0.0001	377.1821	0.6164459	0.9762023
	<i>carotenoids</i>	-1.698561	0.2320054	< 0.0001	417.8862	0.3417714	0.9545229
S1_G	<i>ref. platelets</i>	-2.534945	0.8704142	0.0046	501.861	0.07431972	0.9283915
	<i>melanins</i>	1.471201	1.1052876	0.1868	463.0973	0.01973119	0.9656927
	<i>carotenoids</i>	0.9591593	0.8981691	0.2886	456.3433	0.01206106	0.9450827
S1_Y	<i>ref. platelets</i>	3.867052	2.0905524	0.0679	504.8386	0.03126912	0.9264485
	<i>melanins</i>	-16.35160	1.9005506	< 0.0001	409.8087	0.4420825	0.9722268
	<i>carotenoids</i>	12.520443	1.6114900	< 0.0001	409.7629	0.3719392	0.9558696
S1_R	<i>ref. platelets</i>	2.796298	0.6050707	< 0.0001	491.8553	0.1592639	0.9324678
	<i>melanins</i>	-5.228194	0.5977702	< 0.0001	411.1279	0.4437557	0.9719383
	<i>carotenoids</i>	2.563572	0.6114070	0.0001	441.9675	0.1506411	0.948878

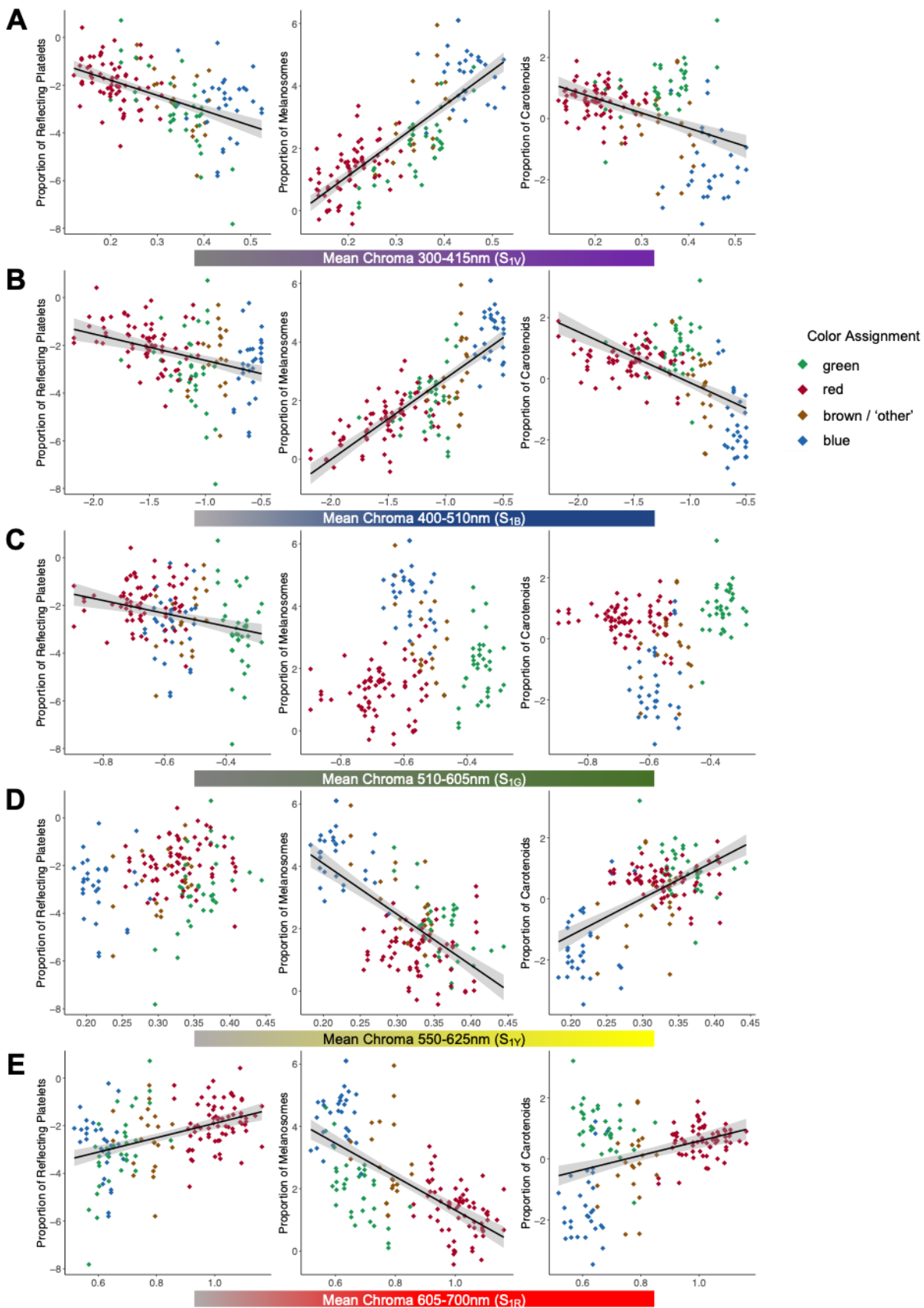


Figure 10. Scatterplot of pivot coordinates of pigment proportions (from left to right: reflecting platelets, melanins, carotenoids) versus chroma metrics (from top to bottom panel: S1_V, S1_B, S1_G, S1_Y, S1_R). Regression lines indicate significant relationships among response and explanatory variables.

2.4.2.3 Dermis, epidermis, and collagen layer heights

I performed a MANOVA to assess the role of dermis, epidermis, and collagen layer heights in explaining color differences. Prior to the MANOVA, I tested model assumptions and screened for outliers. The data do not contain any extreme univariate outliers, but Mahalanobis distance identified three frogs as multivariate outliers, all assigned to color group 2 (green). These individuals were removed from the analysis. Two of 12 color group-variable sets are not normally distributed based on the Shapiro-Wilk test for univariate normality ($p = 0.00250$ for color group 2/epidermis; $p = 0.0268$ for color group 4/epidermis). Data do not exhibit multivariate normality based on the modified Shapiro-Wilk test ($p = 0.000514$). Levene's test of homogeneity of variance was significant ($F = 1.90$, $p = 0.044$) for dermis height, indicating unequal variances, and the Box's M value of 40.8 was significant for color assignment ($p = 0.00164$), indicating heterogeneous variance-covariance matrices. Correlation and linearity among variables appeared suitable for MANOVA, based on linearity plots and correlation values for pairs of dependent variables and each grouping variable (maximum correlation values = 0.598, collagen/dermis/color group 4). To accommodate the numerous assumption violations, I report Pillai's, rather than Wilks', statistic. This statistic is more robust and recommended for unbalanced designs with significant Box's M (Box, 1949).

A two-way repeated-measures MANOVA was analyzed with histosection number (i.e., set 1 or set 2) and color group predicting dermis, epidermis, and collagen heights. Significant

multivariate main effects were found for histosection number ($F_{6,274} = 2.96, p = 0.0081, \eta^2 = 0.122$) and color group ($F_{9,414} = 3.80, p = 0.00013, \eta^2 = 0.229$), but not for their interaction ($F_{18,414} = 0.689, p = 0.822, \eta^2 = 0.087$). Univariate ANOVAs were used to examine individual dependent variable contributions to the main effects. Color groups showed significant differences in dermis height ($F_{3,138} = 8.22, p_{adj} = 0.000406, \eta^2_G = 0.152$) and collagen height ($F_{3,138} = 5.03, p_{adj} = 0.018, \eta^2_G = 0.098$). Histosection number was not a significant effect for either dependent variable.

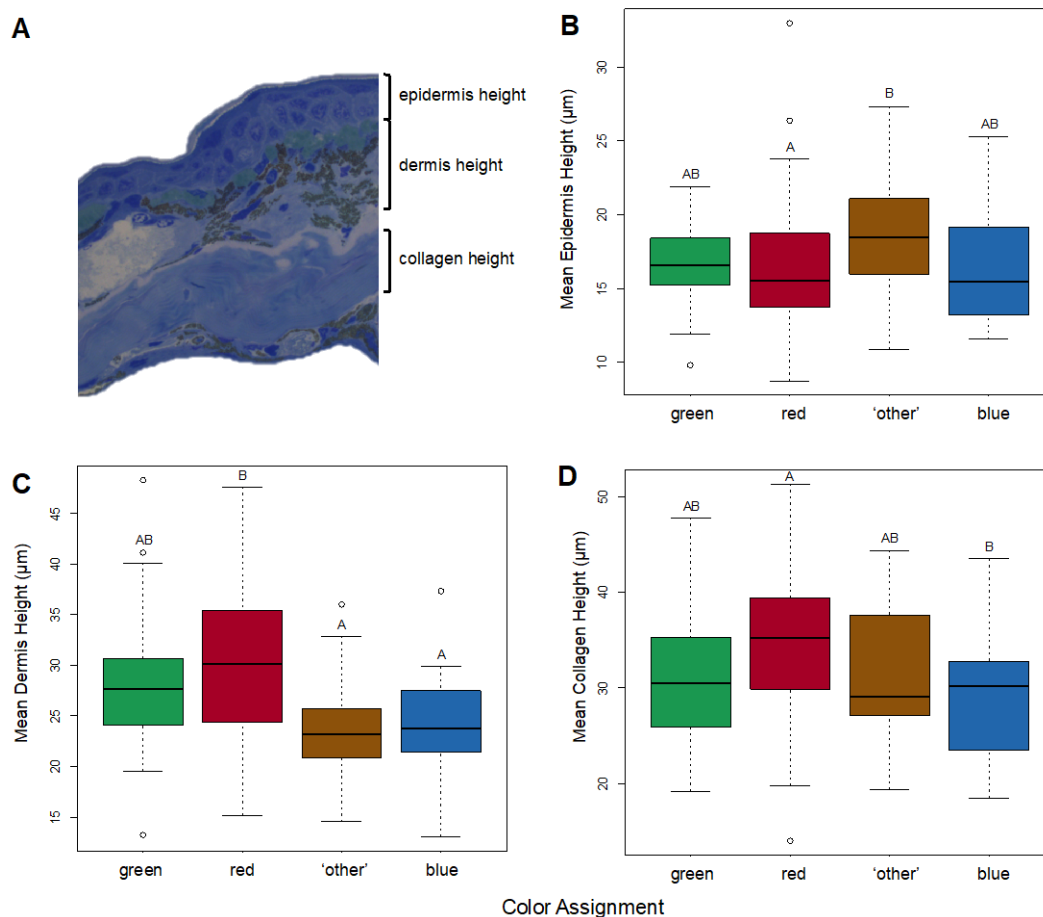


Figure 11. Boxplots showing mean (B) epidermis, (C) dermis and (D) collagen heights, by color group. Significantly different color groups, determined by Tukey post-hoc tests, are indicated by letters. (A) Subset of a histosection showing the location of each measured component in the skin.

I also performed Tukey post-hoc tests to identify pairs of color groups significantly different for collagen and dermis heights (Fig. 11). After Bonferroni corrections, collagen height is significantly different between color groups 2 (red) and 4 (blue; $p_{adj} = 0.00374$), with red frogs having larger collagen layer heights ($m = 34.75 \mu\text{m}$, $sd = 7.44 \mu\text{m}$) than blue frogs ($m = 29.22 \mu\text{m}$, $sd = 6.89$). Dermis height is significantly different between color groups 2 (red) and 3 (brown/'other'; $p_{adj} = 0.000935$) and groups 2 (red) and 4 (blue; $p_{adj} = 0.000349$), with red frogs having larger dermis heights ($m = 30.27 \mu\text{m}$, $sd = 7.95 \mu\text{m}$) than brown/'other' ($m = 23.57 \mu\text{m}$, $sd = 5.35 \mu\text{m}$) and blue frogs ($m = 23.98 \mu\text{m}$, $sd = 5.00 \mu\text{m}$). Epidermis height is significantly different between groups 2 (red) and 3 (brown/'other'; $p_{adj} = 0.0269$), with red frogs having smaller heights ($m = 16.24 \mu\text{m}$, $sd = 4.19 \mu\text{m}$) than brown/'other' frogs ($m = 19.03 \mu\text{m}$, $sd = 3.90 \mu\text{m}$).

2.4.2.4 Pigment proportions and layer heights as predictors of color groups

I compared LDA models with and without height data to investigate how well pigment proportions or epidermis, dermis, and collagen layer heights separate individuals by color. After testing assumptions and screening for outliers, I removed one univariate and one multivariate outlier from the analysis. LDA results are summarized in tables 4 and 5.

Table 4. Coefficients of linear discriminants for LDA with and without height data (with/without).

Loading magnitude indicates its contribution to separating individuals among color groups, relative to other variables. Negative values indicate the discriminant function score for that variable is predicted to decrease when other variables are held constant.

	LD1 with / without height	LD2 with / without height	LD3 with / without height
% separation achieved	92% / 96%	5% / 4%	3% / na
reflecting platelet	0.00 / -0.03	-0.63 / -0.87	-0.43 / na
melanin	-1.34 / 1.36	-0.34 / -0.21	0.10 / na
dermis height	0.68 / na	-1.55 / na	4.79 / na
epidermis height	0.27 / na	2.44 / na	-5.19 / na
collagen height	0.55 / na	-1.96 / na	-2.48 / na

Table 5. Prediction accuracy for resubstitution and leave-one-out cross validation methods.

	Resubstitution		Leave-one-out Cross Validation			
	<i>Overall</i>	<i>Overall</i>	<i>Red Group</i>	<i>Green Group</i>	<i>Other / 'brown' Group</i>	<i>Blue Group</i>
pigment proportion only	0.68	0.68	0.31	0.91	0.19	0.87
pigment proportion + heights	0.68	0.66	0.41	0.85	0.19	0.83

The apparent error rates and linear decision boundaries for both models are very similar (pigment proportions only = 0.325, pigment proportions + heights = 0.311), suggesting that including layer heights only marginally improved the model's ability to separate of samples by color group. Stacked histograms of the first discriminant function value for samples from different color groups further illustrate the minimal effect of height data on separating samples by color (Fig. 12). The first discriminant function in both models clearly separates green and red frogs from group 4 blue frogs.

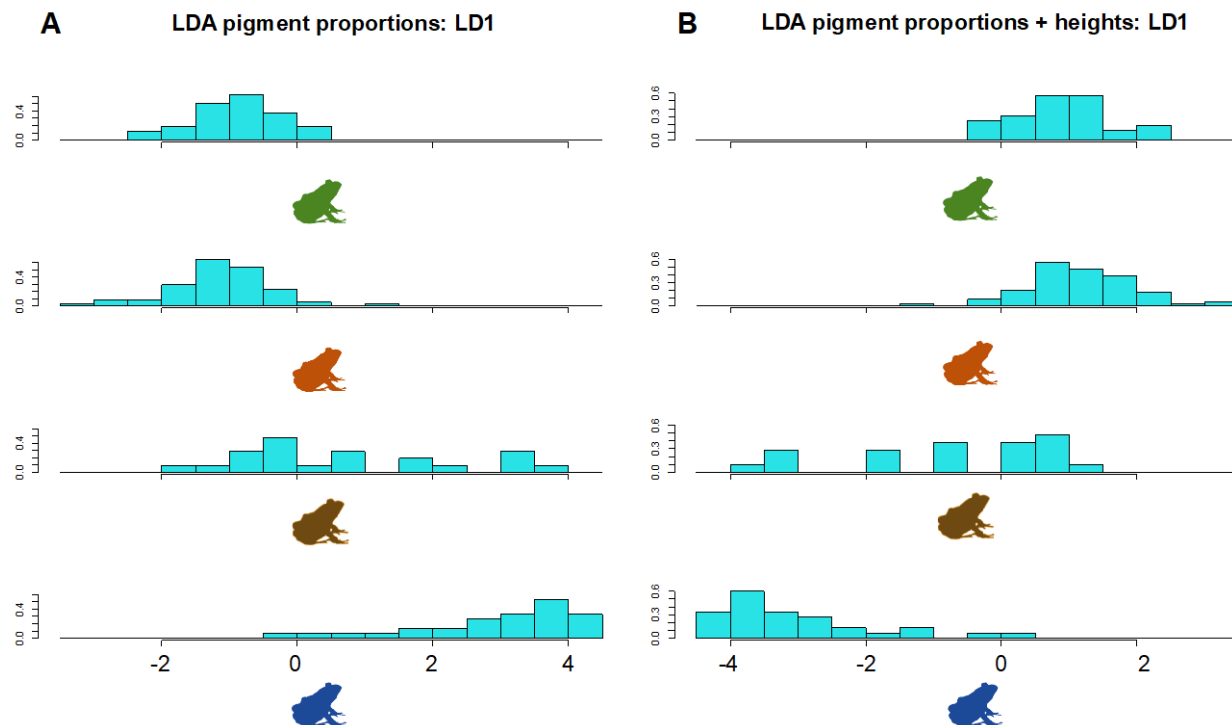


Figure 12. Stacked histograms of the first discriminant function value for samples from different color groups.

Histograms for (A) LDA of pigment proportion data only and (B) LDA of pigment proportions plus collagen, epidermis, and dermis heights. In both analyses, the first discriminant function clearly separates color green and red frogs from blue frogs. Including heights in the model does not appear to improve classification.

Table 4 lists each variable's linear coefficients (loadings) and the percent of among-group variation explained by the discriminant functions. For LDAs with and without height data, melanin has the strongest loading on the first discriminant function, indicating this pigment contributes most to its separation. In the LDA with height data, epidermis height has the strongest loading on the second discriminant function, with dermis and collagen heights showing a slightly smaller contribution. The reflecting platelet pivot coordinate has weak loadings on LD1 and LD2 for models with and without height data.

The prediction accuracy using the resubstitution error is the same for models with and without height data (Table 5). However, discriminant analyses performed directly on observed data, without training data, generally underestimate the apparent error rate. To this end, I repeated the LDA with leave-one-out cross-validation, but model accuracy did not significantly change with this approach. In both models, prediction accuracy was highest for color groups 2 and 4 (Table 5) and lowest for group 3; accuracy for group 1 increased from 31% to 41% with the inclusion of heights data.

2.4.2.5 Chemical determinants and differences among color groups

The HPLC and UV/VIS spectroscopy analyses produced two main findings. First, there is no evidence of pterins in any of the *O. pumilio* skins. Second, HPLC analysis revealed an extreme diversity of carotenoid pigments, greater than many other vertebrates analyzed by the McGraw lab (McGraw, *pers. comm.*). There are 17 identified carotenoid species, and 2 unknowns (Appendix Table 6). Non-zero concentrations range from 0.08 $\mu\text{g/g}$ (canary xanthophyll, NI, Isla Pastores) to 335.27 $\mu\text{g/g}$ (β -carotene, FA, mainland). To parse this unprecedented carotenoid diversity, I grouped esterified pigments and cis-isomers with their unmodified forms, then calculated the percentage of each dorsal carotenoid by color group. The most apparent finding is how few carotenoids are present in blue frogs; just two types of carotenoid pigments are found in blue frogs (apocarotenoids, canary xanthophylls), compared to 9 major types (apocarotenoids, β -carotene, canary xanthophylls, canthaxanthins, echinenones, ketocarotenoids, luteins, and xanthophylls) found in the yellow and red frogs.

I applied non-metric multidimensional scaling (NMDS) to HPLC abundance data from dorsal *O. pumilio* skins to visualize carotenoid suites' dissimilarities by color group (NMDS ordination stress = 0.06). The resulting NMDS plot (Fig. 13) illustrates a carotenoid composition

distinction based on color group. In this plot, red morphs, and the sole yellow morph (CY), are dissimilar from blue/brown and green morphs.

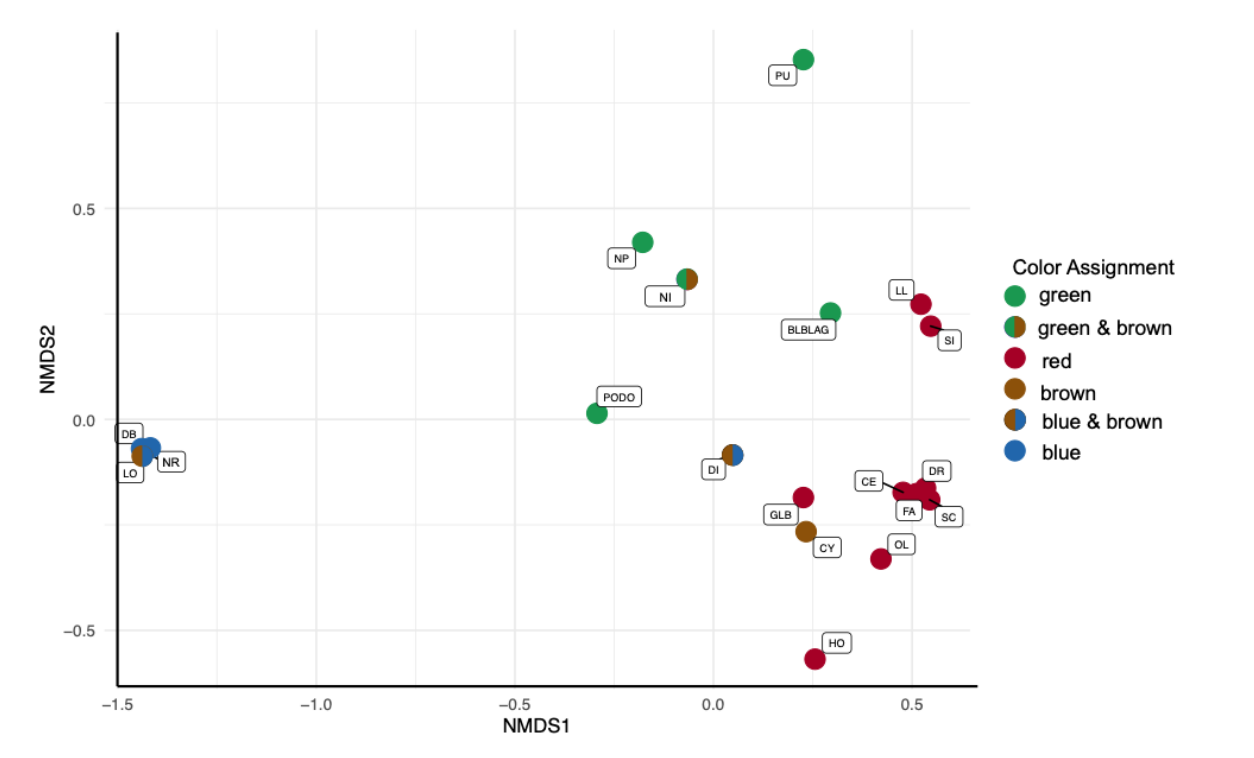


Figure 13. NMDS plot of HPLC abundance data from dorsal *O. pumilio* skin, illustrating differences in carotenoid composition between some *mclust*-derived color groups.


Data points are labelled with population/morph abbreviations (Fig. 6). Populations containing individuals assigned to more than one color group (e.g., green & brown, blue & brown) are labelled accordingly. Red morphs are dissimilar from blue/brown and green morphs.

ANOSIM suggests statistically significant differences in carotenoid suites among color groups ($R = 0.704$, $p = 0.0001$). SIMPER tests show 10 carotenoid species are found significantly more often in one color group than another (Table 6, Appendix Table 6). Green and red color groups show significantly different abundances of β -carotene and cis lutein ester, hinting at a potential mechanism for color differences in these groups. There are no significant differences in carotenoids between green and brown groups or red and brown groups. The greatest number of

significant pairwise contrasts include blue frogs, an unsurprising result given their more uniform carotenoid suites and low carotenoid abundances (Appendix Figure 1), relative to other colors. However, separate suites of pigments explain differences between blue and red frogs (canthaxanthin ester 1, canthaxanthin ester 2, and cis lutein ester) and blue and green frogs (apocarotenoid, canary xanthophyll esters 1, 2, and 3, and ketocarotenoid ester 2).

Table 6. A summary of significant SIMPER results.

Blue vs. green, brown, and red frogs show a significant difference in carotenoids, but separate suites of pigments explain their differences. Red and green frogs, which are characterized by similar proportions of pigments, have different amounts of two carotenoids, implicating carotenoid processing as a potential explanation to explain their color difference.



apocarotenoid** canary xanthophyll ester 1*** canary xanthophyll ester 2*** canary xanthophyll ester 3* ketocarotenoid ester 2**	echinenone*** canthaxanthin ester 1*** canthaxanthin ester 2*	β -carotene* canthaxanthin ester 1* cis lutein ester*	β -carotene* cis lutein ester*
--	---	---	---

Significance codes: 0 '***' 0.001 '**' 0.01 '*'

Since I derived the color groups for ANOSIM and SIMPER tests using the first two principal components of chroma metrics (see *mclust* methods, above), I performed Mantel tests to determine the correlation between carotenoid suites and B_I , S_R , S_Y , S_G , S_B , and S_V . I calculated dissimilarity matrices using Bray-Curtis dissimilarity for carotenoid data and Euclidean distance for the colorimetrics. Mantel tests, summarized in Table 7, indicate significant correlations between differences in carotenoid suites and color variables for all but the S_G metric.

Table 7. Summary of MANTEL tests to estimate the correlation between carotenoid suites and individual color metrics.

The color metrics include total brightness (B_I) and the five chroma metrics described in Fig. 7.

Differences in carotenoid suites are correlated with differences in chroma metrics for every metric except for S_{IG}.

	S metrics combined	S _{IR}	S _{IY}	S _{IG}	S _{IB}	S _{IV}	B _I
Mantel r statistic	0.6438	0.47	0.5059	0.03036	0.6376	0.6775	0.3033
<i>p</i> value	0.0001	0.0003	0.0003	0.3348	0.0001	0.0001	0.0088

To investigate the relationship between individual carotenoid pigments with total brightness and chroma metrics, I performed linear models with the color metric variable as the response and carotenoid pigment as the predictor. The results of these models are summarized in Table 8 and Appendix Table 7. Several carotenoid pigments predict increases in total brightness, including canary xanthophyll esters 1 and 2, β -carotene, ketocarotenoid ester 2, cis lutein ester, and lutein ester. Chroma in the red wavelengths show a positive relationship with all but one pigment, for which the effect is non-significant. In contrast, yellow and green wavelengths show a significant relationship with only one pigment (respectively: ketocarotenoid, positive; apocarotenoid, negative). Chroma in the blue and violet ranges share near-identical significant negative relationships with canary xanthophyll esters 1, 2, and 3, β -carotene, ketocarotenoid ester 2, cis lutein ester, and lutein ester 1. Among all the pigments tested with linear models, ketocarotenoid ester 2 shows a significant relationship across the broadest range of wavelengths (Table 8). This pigment's shift from negative to a positive association, across long to short wavelengths, closely mirrors the relationship modeled with carotenoid proportion and chroma

metrics (Fig. 10A-E, center column). The same pattern is evident for all the carotenoid pigment-chroma metrics modeled.

Table 8. Summary of linear models of the relationship between individual carotenoid pigments and colorimetric variable.

Positive (+) or negative (-) symbols indicate direction of the linear relationship; significance codes indicate the level of significance for each model. The color metrics include total brightness (B_I) and the five chroma metrics described in Fig. 7. Chroma in the red wavelengths show a positive relationship with all but one pigment.

Carotenoid Pigment	B _I	S _{1R}	S _{1Y}	S _{1G}	S _{1B}	S _{1v}
Apocarotenoid	ns	+ **	ns	- ***	ns	ns
Canary Xanthophyll Ester 1	+ *	+ **	ns	ns	- *	- *
Canary Xanthophyll Ester 2	+ *	+ **	ns	ns	- *	- *
Canary Xanthophyll Ester 2	ns	+ **	ns	ns	- *	- *
Echinenone	ns	ns	ns	ns	ns	ns
Canthaxanthin Ester 1	ns	+ *	ns	ns	ns	ns
Canthaxanthin Ester 2	ns	+ *	ns	ns	ns	ns
β-carotene	+ ***	+ ***	ns	ns	- **	- **
Ketocarotenoid Ester 2	+ *	+ *	+ *	ns	- *	- *
Cis Lutein Ester	+ ***	+ ***	ns	ns	- **	- **
Lutein Ester 1	+ **	+ **	ns	ns	- *	- *

Significance codes: 0 '***' 0.001 '**' 0.01 '*' not significant 'ns'

Lastly, to assess agreement between morphological and HPLC datasets, I modeled the relationship between the log-transformed total carotenoids from HPLC and (1) total brightness, B_I and (2) the proportion of carotenoids I measured from histological sections. Since there is not perfect overlap between histological and HPLC datasets, I removed populations with missing data (patterned collecting sites in Figure 6). The results of the first linear regression indicate total carotenoids from HPLC explain 41.3% of the variance ($R^2_{adj}=.413$, $F_{1,16}=12.97$, $p=0024$) and

predict total brightness ($t = 6.02, p = 0.00239$; Appendix Figure 1). This result is in agreement with findings from histological data (Fig. 9B), suggesting congruence between the two datasets. The results of the linear regression with histological data indicate total carotenoids from HPLC explain 41.1% of the variation ($R^2_{\text{adj}} = 0.411, F_{1,16} = 12.86, p = 0.0025$) and predict carotenoid proportions measured from histological sections ($t = 3.59, p = 0.00247$; Appendix Figure 2). The scatterplot suggests that my histological measurements of carotenoids are more accurate for red frogs than blue or green frogs.

2.5 Discussion

Taken together, the results suggest that large color variations in *O. pumilio* are explained by differences in pigment proportions and carotenoid suites. I detected significant differences in the dermis, epidermis, and collagen layer heights between some color groups, but these components distinguish individuals by color to a much lesser extent than pigment proportions. The total proportion of pigmented dermis is not correlated with chroma or brightness and does not appear to differ among color groups (Appendix Table 5). In all but one population sampled, all three pigment cells were present, suggesting that the loss of pigment cell types is not a mechanism by which color change in this species commonly occurs.

Several results from my analyses deserve further discussion. First, the proximate mechanism of blue coloration in poison frogs, as revealed by histology and HPLC data, differs from previous studies. I found increases in violet and blue chroma metrics are associated with decreases in proportions of reflecting platelets (Fig. 10, Table 3). Traditionally, amphibians' blue coloration is attributed to the reflecting and scattering properties of reflecting platelets (Bagnara

et al., 2007). Moreover, a previous description of a blue *O. pumilio* morph, based on a micrograph from transmission electron microscopy (TEM), reports multiple iridophore layers and an absence of xanthophores (Duellman & Trueb, 1986). Twomey et al. (2020) reach the same conclusion from a single sample of blue *Dendrobates tinctorius* they analyzed by TEM, despite finding chemical evidence of pterins. In this study, histosections of some blue individuals lack xanthophores, but I ultimately found evidence of carotenoids in all but one blue morph. This was supported by the HPLC analyses, which returned non-zero concentrations of carotenoids for two blue morphs. The disparity between the aforementioned and present studies underscore TEM's limitations and the importance of examining multiple histosections from multiple individuals of each morph.

Second, few studies of the morphological basis of coloration consider components other than chromatophores. The collagen layer has documented reflective properties in frogs (Nielsen & Dyck, 1978) and is even incorporated into models for the generation of color (Bagnara & Hadley, 1973; Nielsen & Dyck, 1978). As summarized in Grether et al. (2004), a reflective collagen layer can return through the DCU any light unabsorbed by the melanophore layer. As such, when a reflective collagen layer is present, the amount of melanin will influence how much light passes through the absorptive xanthophore layer for a second time (Grether et al., 2004). Despite this knowledge, many microscopy analyses fail to incorporate the collagen layer. This study found that red *O. pumilio* individuals have significantly greater collagen layer heights than blue frogs (Fig. 11). As shown in the scatterplots for melanin proportion against chroma metrics (Fig. 10A-E, center column), red frogs generally have lower melanin proportions than blue frogs. If a thicker collagen layer indicates greater reflective properties, then the integument of red frogs is modified in at least these two ways to increase the xanthophore layer's influence. Although this study also

found significant differences in the epidermis and dermis heights between some color groups, their effects, if any, on color production are not understood.

Third, the diversity of carotenoid pigments in *O. pumilio*, and the different compositions between color groups, highlight the important role of carotenoid coloration in this system. Red frogs show significantly greater β -carotene and cis lutein ester abundances than green frogs. However, different pigments contrast blue and red frogs (canthaxanthin ester 1, canthaxanthin ester 2, and cis lutein ester) and blue and green frogs (apocarotenoid, canary xanthophyll esters 1, 2, and 3, and ketocarotenoid ester 2; Table 6). Though I did not detect significant differences between red and yellow frogs' carotenoid suites, this may be an issue of statistical power ($N_{\text{red}} = 9$; $N_{\text{yellow}} = 1$; Appendix Table 6). Recently, Twomey et al. (2020b) used thin layer chromatography (TLC) to examine carotenoid profiles in red and yellow *Ranitomeya sirensis* morphs. The ketocarotenoids echinenone, 3-hydroxyechinenone, canthaxanthin, adonirubin, and astaxanthin were present in red frogs but absent in the yellow morph, a polymorphism they link to differential expression of carotenoid processing genes. It is interesting that red/yellow *O. pumilio* morphs from the polymorphic Cemetery Hill population do not show these differences; almost all of the pigments present in red frogs are also present in yellow frogs, albeit in lower amounts (not shown). It is possible that entirely different molecular mechanisms may be responsible for the yellow/red polymorphism in *O. pumilio*, a hypothesis that future studies in this system should address.

Finally, this study suggests that *O. pumilio* lacks pterins, a finding that disagrees with the literature. Thin-layer chromatography (TLC) from a sample of *O. pumilio* showed trace quantities of the pteridine drosoplerin (Twomey et al., 2020). A differential expression study in *O. pumilio* (Rodriguez et al., 2020) identified eight genes related to pteridine synthesis that are more

highly expressed in green and red frogs. Given these findings, and the prevalence of pterins in other poison frogs (Posso-Terranova & Andrés, 2017; Twomey et al., 2020), I suspect the assay used in this study is not sufficiently sensitive to detect small amounts of pterins.

Studies of coloration in various animal systems have provided insights into the genetic basis of convergent evolution. Reviews of the proximate and ultimate causes of melanin-based color convergence are numerous and underscore its highly-conserved genetic underpinning among vertebrates (Hoekstra, 2006; Manceau et al., 2010). Among lizards from White Sands, New Mexico, rapid convergence on white phenotypes is attributed to mutations in the same gene (melanocortin-1 receptor gene, *Mclr*), but the molecular mechanisms that result in less melanin are different (Rosemlum et al., 2010). McRobie et al. (2019) arrived at similar results from studying the *Mclr* gene in two tree squirrel species, where independent genetic mechanisms are responsible for convergence on melanistic patterning.

The proximate mechanisms of convergence in bright, aposematic, and mimetic phenotypes are far less studied than melanin-based coloration, likely due to the DCU's underlying complexity. Nevertheless, studies across poikilotherm systems have addressed this disparity (Kikuchi et al., 2014 (snakes), Prager & Andersson, 2010; Prum & Torres, 2003; Twyman et al., 2017 (birds), Dick et al., 2018 (fish), Twomey et al., 2020; Posso-Terranova & Andrés, 2017 (frogs)). Avian taxa, in particular, have emerged as popular subjects for understanding the genetic basis of carotenoid metabolism and modification (Toews et al., 2017). Scavenger receptors, β -carotene oxygenases, and ketolases, representing several classes of genes, are implicated in the transport, deposition, and metabolic processing of carotenoids in birds (Toews et al., 2017). Interestingly, many of these genes are conserved across deep phylogenetic distance (e.g., birds and turtles).

Recent studies in poison frogs have begun to disentangle the complexity of color production mechanisms in dendrobatids, and in aposematic amphibians more generally. Twomey et al. (2020) used TLC, absorbance spectrometry, and transmission electron microscopy to examine the ultrastructural basis of coloration in the *Ranitomeya imitator* mimicry complex. Models with these data as parameters found reflecting platelet thickness and the pteridine drosopterin, but not carotenoids, explain coloration in *R. imitator* and its model species. In contrast, Posso-Terranova and Andrés (2017) suggest harlequin poison frog (*O. histrionica*) color differences are explained by xanthophores and melanosome dispersion. In both studies, sample sizes for microscopy data are fewer than 3 individuals per species/morph, and may explain some of the discrepancies between their conclusions.

Before the present study, the cellular, pigmentary, and structural basis of coloration in *O. pumilio* were largely unknown. However, phylogeographic, genomic, ecological, and behavioral studies in the species promise a fascinating system to study the genetic basis of color variation and convergent evolution. Character reconstructions of dorsal coloration in *O. pumilio* suggest multiple changes from bright (yellow, red, orange) to dull (green and brown) coloration (Wang & Shaffer, 2008). Summers et al. (2003) speculate reversals to green coloration could be due to convergence on pigmentation, but offer ancestrally retained synthesis pathways as the more likely explanation.

Although this study represents the early stages of understanding the proximate mechanisms of coloration in *O. pumilio*, the results motivate interesting future research topics. My results indicate that shifts from red to green coloration are accomplished by decreasing reflecting platelets (Fig. 10) and mutations that suppress *cis* lutein ester or carotene in the skin. Additional candidate gene and transcriptome studies in *O. pumilio* can target these and other morphs with divergent

carotenoid suites (Table 6). Larger sampling sizes, especially among red morphs, may reveal differences in carotenoid suites and the DCU where this study did not. This would allow testing hypotheses *within* color groups to examine the relative roles of xanthophores, iridophores, and melanophores in producing narrow spectral variation. Finally, to provide a complete picture of color-producing mechanisms, future studies should integrate the developmental basis of coloration with morphological and genetic analyses.

3.0 Estimating the Phylogenetic Relationships Among Closely Related Populations of *Oophaga pumilio* from Bocas del Toro, Panama

3.1 Summary

Oophaga pumilio has emerged as a popular species for evolutionary biologists interested in the roles of geographic isolation and selection in driving phenotypic divergence. Although this species provides a compelling example of phenotypic variation and incipient speciation, the system is characterized by many of the challenges associated with studying recently separated populations. In this study, I provide a high-resolution species tree for *O. pumilio* by incorporating a more focused sampling strategy (greater number of individuals, greater number of collecting sites), the application of NGS, and population genetic analyses. To test whether dorsal coloration is the result of convergent evolution, I use the topologies and novel spectral reflectance data to perform ancestral character state reconstruction. The topological relationships evident in the phylogenies and the distribution of dorsal color phenotypes suggest that vicariance largely explains genetic differentiation pattern in *O. pumilio*. The ancestral character state reconstructions support the convergent evolution of dull colors like green and blue and at least one reversal to bright coloration like yellow, orange or red. Incorporating these species trees with future genetic, morphological, and physiological datasets will provide an integrative approach to study color polymorphism as a driver of speciation.

3.2 Introduction

The strawberry poison frog, *Oophaga pumilio*, is a fascinating example of color polymorphism. *Oophaga pumilio* is abundant throughout the lowlands on the Caribbean slopes of Panama, Costa Rica, and Nicaragua, where populations are typically composed of a single, aposematic color morph characterized by a bright red body with blue/black limbs (Daly & Myers, 1967; Summers et al., 2003). However, on the Bocas del Toro archipelago of western Panama and the adjacent mainland, *O. pumilio* exhibits some of the most striking variation in color and pattern seen in nature, with numerous color morphs described (Hoffman & Blouin, 2000). Throughout the archipelago region, island and adjacent mainland populations possess distinct phenotypes, with each population generally consisting of a single color morph (Summers et al., 2003). Exceptions to this, where multiple color morphs occur in sympatry, are found on the Aguacate Peninsula, located on the mainland adjacent to the islands, and in populations on Isla Bastimentos. These polymorphic populations consist of red, brown, and blue frogs (Aguacate Peninsula) and bright yellow and red frogs with black spots (Isla Bastimentos).

The divergence of the numerous, highly variable color morphs of *O. pumilio* from their hypothesized bright red ancestor, as suggested by numerous phylogenetic studies (Summers et al., 1997; Hagemann & Pröhl, 2007; Wang & Shaffer, 2008; Hauswaldt et al., 2010) and geographic data (Anderson & Handley, 2002; Coates et al. 2005), occurred approximately 10,000-1,000 years ago during the formation of the Bocas del Toro archipelago. Rising sea levels in the region first separated Isla Colón (~6,300 years ago), then the land that makes up present-day Isla Bastimentos and Isla Solarte (~4,700 years ago), followed by the separation of Isla Cayo Agua (~3,400 years ago). The islands reached the current configuration approximately 1,000 years ago, when Isla Popa and Isla San Cristobal separated from the mainland, and Isla Solarte split from Bastimentos

(Anderson & Hadley, 2002). The sequential, recent, and rapid formation of the Bocas del Toro archipelago provides for a uniquely challenging backdrop to study the co-distribution of genetic and phenotypic diversity, and the relative roles of selective and neutral processes during phenotypic evolution.

The potential complexity of interactions between natural and sexual selection on coloration has been illustrated by molecular, chemical, and behavioral studies of *O. pumilio*, recently summarized by Gehara et al. (2013). Unlike many other aposematic species, *O. pumilio* exhibits extreme color and pattern variation, despite the theoretical predictions that individuals of an aposematic species will exhibit a limited number of phenotypes, so as to maximize the effectiveness of their warning signal (Müller, 1879). Furthermore, females of this species exhibit color-assortative mating in the wild (Richards-Zawacki et al., 2012; Yang et al., 2019) and in laboratory experiments (Reynolds & Fitzpatrick, 2007; Mann & Cummings, 2008), suggesting the possibility for sexual selection against novel phenotypes in wild populations. However, the effect of assortative mate preferences on color evolution is mediated by male-male competition, suggesting the complexities of phenotypic evolution (Yang & Richards-Zawacki, 2020). Together, these features of *O. pumilio* result in a system that provides an excellent opportunity to study how natural selection, sexual selection, and geographic isolation on islands- and the interactions among these processes- contribute to the evolution and maintenance of an ecologically important phenotypic trait such as coloration.

The evolution of color variation in the strawberry poison frog, *Oophaga pumilio*, has been addressed by a number of studies utilizing combinations of population genetics, phylogenetics, phylogeography, and phenotypic analyses. The first such study, published 17 years ago by Summers et al. (1997), compared levels of mtDNA sequence divergence between *O. pumilio* and

two sympatric species of frogs (*Phyllobates lugubris* and *Minyobates* sp.) that exhibit little phenotypic variation across the Bocas del Toro archipelago. Based on their cladistic and distance analyses of 292 bp of mtDNA cytochrome *b*, the authors found similarly low levels of sequence divergence in all three species. These results prompted the authors of this early study to reject a neutral model of evolution in which genetic drift coupled with longer periods of population isolation leads to greater interpopulation mtDNA variation and subsequently greater phenotypic divergence in *O. pumilio*. Instead the authors support a model of sexual selection associated with female parental care. In this model, unequal investment between males and females generates the potential for strong selection, allowing for the rapid evolution of male ornamentation that is driven by female preference.

Similar conclusions were drawn from a larger phylogenetic study that included samples ranging from northern Costa Rica to the Bocas del Toro archipelago and used mtDNA sequences from three genes (cytochrome *b*, cytochrome oxidase I, and 16S rRNA; Hagemann and Prohl, 2007). The authors found larger genetic differences among monomorphic Costa Rican populations than among polymorphic Bocas del Toro populations, suggesting that the independent evolution of island populations since the formation of the archipelago cannot account for phenotypic differences among island populations. A thorough phylogenetic investigation of these data using maximum likelihood, maximum parsimony, and Bayesian approaches recovered three distinct monophyletic groups within *O. pumilio*. More importantly, the inclusion of homologous mtDNA from related species revealed the genetic paraphyly of *O. pumilio*, that is, some monophyletic groups within the *O. pumilio* phylogeny were more similar to related species than they were to other monophyletic groupings of their own species. Although both of the aforementioned studies reject a model of neutral evolution, Hagemann and Prohl (2007) invoke differences in both natural

and sexual selection pressures as the likely cause of color polymorphism in *O. pumilio*, while Summers et al. (1997) considers only sexual selection. Rudh et al. (2007), using amplified fragment length polymorphisms (AFLPs) in combination with morphological data from snout vent length (SVL), spot patterns, and coloration, conclude that natural selection, sexual selection, and neutral processes may be involved in the evolution of color variation in *O. pumilio*.

Wang and Shaffer (2008) analyzed 3051 bp of mitochondrial DNA (mtDNA) sequence data and reconstructed ancestral character states using color data obtained from photographs. Ancestral character state reconstruction, whereby morphological characters are coded and mapped over phylogenetic hypotheses, is a potentially powerful tool used to elucidate the evolutionary history of phenotypic traits (Schluter et al., 1997). Using this method, Wang and Shaffer (2008) found evidence for the maximal number of changes in all color traits, suggesting that convergence has driven the repeated evolution of bright coloration as well as reversals to dull dorsal coloration, invoking roles of both natural and sexual selection. The authors also reject a model of strict vicariance due to increased isolation during island formation. Three more recent studies utilized mtDNA (Brown et al., 2010), a combination of mtDNA, nuclear DNA, and microsatellites (Hauswaldt et al., 2010), and solely microsatellites (Wang and Summers, 2010). Each of these studies included more sequences and denser sampling than previous work. Their results all highlight the role of selection in generating variation in color, yet none of these studies have included more than a few samples per island and all have based their inferences on information gleaned from a small portion of the genome. Furthermore, while all of the studies described above, from the pioneering study of Summers et al. (1997) to more recent works, have provided important contributions to understanding the extreme variety of color patterns in *O. pumilio*, the phylogenetic relationships among closely related populations are not fully understood. Moreover, the

phylogenies resulting from these studies are characterized by a lack of reciprocal monophyly and often contradict one another in ways that preclude inferences of how populations have evolved.

Next-generation sequencing (NGS) allows the processing of millions of DNA sequences in parallel, thereby permitting genome-wide scans in a cost- and labor- efficient manner. One particular NGS method, developed by Baird *et al.* (2008), uses restriction enzyme cut sites to anchor assays of genetic variation. This method, known as Restriction site Associated DNA (RAD) markers, generates genome-wide sequence tags (RAD tags) that can be used to identify thousands of single nucleotide polymorphisms (SNPs) across individuals and populations. This method has proven to be a useful application of NGS to phylogenetics and phylogeography of countless organisms.

A high-resolution estimate of the pattern of genetic divergence among populations is needed to provide the framework to explore the evolutionary processes that have generated the remarkable variation in coloration and pattern that we see in *O. pumilio*. In this study, I aim to infer a high-resolution species tree of recently diverged populations from Bocas del Toro by utilizing RADseq. To test whether dorsal coloration is the result of convergent evolution, I use the topologies and novel spectral reflectance data to perform ancestral character state reconstruction.

3.3 Methods

3.3.1 Sampling Details

I collected 10-20 *O. pumilio* individuals from each of 22 locations throughout the Bocas del Toro region of Panama (Fig. 14, Appendix Table 8). Rather than limit sampling to populations

that vary in coloration, I chose localities so as to bracket hypothesized breaking points during island formation (Anderson & Handley, 2002; dashed lines in Fig. 14) while also including most of the known phenotypic variants. The ancestral Costa Rican morph was represented by four individuals from La Selva and Puerto Viejo.

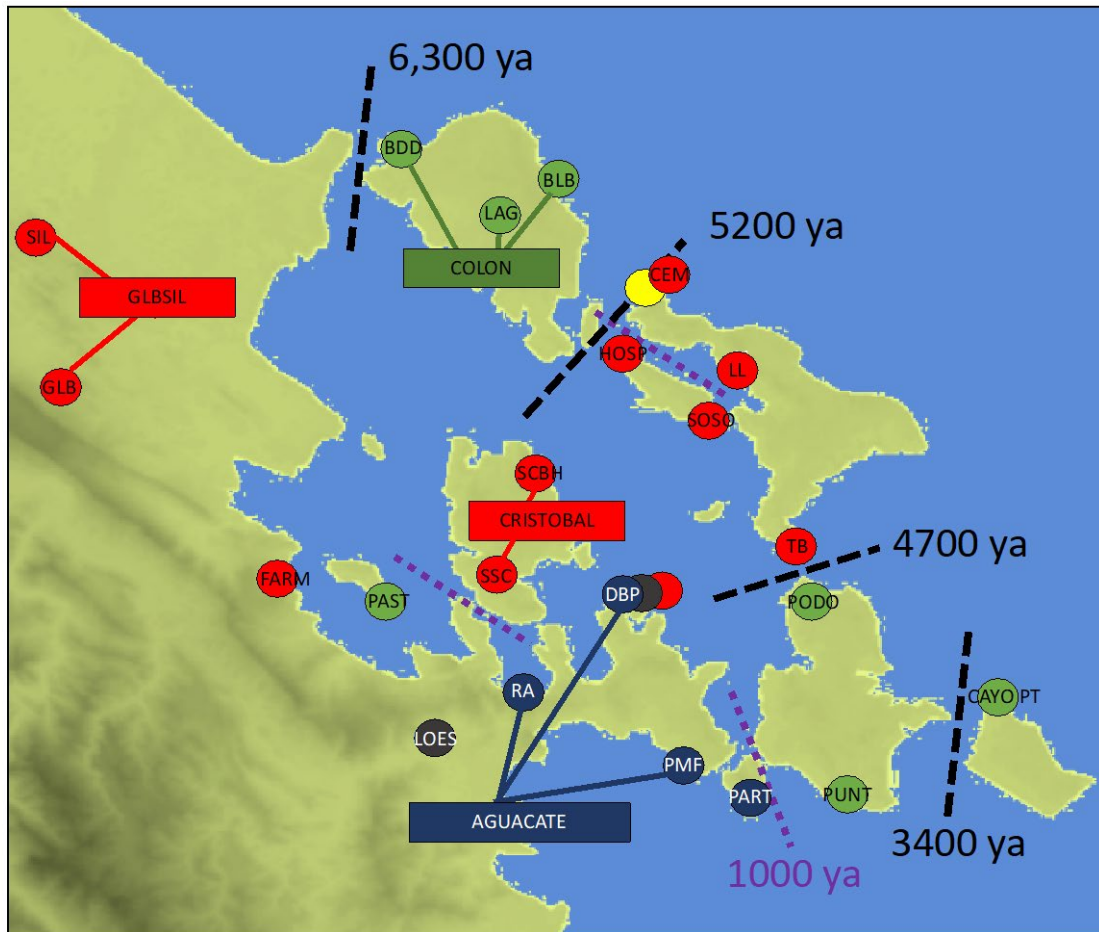


Figure 14. Map of Bocas del Toro, Panama, showing the location of samples included in this analysis. Points are colored according to by-eye dorsal coloration of frogs at each location; overlapping points indicate polymorphic populations. Collecting sites that were ultimately assigned to the same population are indicated by solid lines. The age of geographic separation between islands, adapted from Wang & Shaffer (2008), is indicated by dashed lines. The most recent splits, occurring within the last 1000 years, are indicated by purple dashed lines.

3.3.2 ddRADseq: Library Preparation, Sequencing, and Bioinformatic Processing

I followed a double-digest RADseq protocol (DaCosta & Sorenson, 2014) with restriction enzymes EcoRI and SbfI and sequenced one, pooled library, containing individual-level barcodes, on two Illumina lanes. Raw reads (average = 1,097,264 per individual) were demultiplexed and then trimmed with *Trimmomatic V0.32*). Sequences present at a frequency > 1% of total reads for any individual were removed, resulting in an average of ~400,000 reads per individual. Reads were aligned to the *O. pumilio* genome (Rogers et al., 2018) with *BWA-MEM* and filtered for mapping quality > 30 using *samtools* (average read survival = 65%). SNPs were identified with *freebayes* variant detector, excluding indels and limiting the number of alleles considered to four.

I followed a stringent variant filtering scheme for RADseq data adapted from O’Leary et al. (2018). First, I removed low-confidence SNP calls (minQ 30, minDP 10, mac 3) and applied low cut-off values for missing data (individuals and loci), followed by iteratively increasing cut-off values to individual missing data > 25% and loci call rate < 95%. The final, filtered dataset contains 2,489 biallelic SNPs on 673 scaffolds from 235 individuals. The filtering steps, with the number of SNPs and individuals remaining at each, are outlined in Appendix Figure 3.

3.3.3 Population Differentiation

Genome-wide patterns of divergence among the 22 *O. pumilio* populations were assessed using fineRADstructure (Malinsky et al., 2016). This program infers population structure from haplotype linkage information and a coancestry matrix of haplotype relationships, allowing for the inference of recent or current population structure from RADseq data. It has been successfully

used in many systems confounded by recent, shared ancestry (Stryjewski & Sorenson, 2017; Peart et al., 2017; Lavretsky et al., 2019). Malinsky et al. (2016) suggest users caution against large, systematic differences in missing alleles among putative populations, as it can lead to spurious results. To address this concern, I visually inspected missingness (the proportion of missing alleles) among collecting sites by plotting the ‘missingness’ output from fineRADstructure.

3.3.4 Phylogenetic Analysis

I estimated species trees from SNP data with individuals assigned to populations based on the coancestry matrix. I employed two multi-species coalescent analyses (MSC) appropriate for RAD loci. The first analysis, SNAPP, is implemented in BEAST and assumes that shared alleles between lineages are due to incomplete lineage sorting and not gene flow (Bryant et al., 2012). SNAPP also assumes unlinked, bi-allelic loci, so I selected one SNP per scaffold. I tested the remaining SNPs for linkage, then removed one per linked pair with the package *GBS SNP Filter* in R. SNAPP run times increase exponentially with the number of operational taxonomic units (OTUs), so it can be prohibitively slow for some analyses. To this end, I ran two separate analyses with 5 non-overlapping individuals per OTU. I used BEAUTi to generate the input file with the following parameters: $u = 1.0$, $v = 1.0$, coalescent rate = 10.0, default lambda prior, $\alpha = 1$, $\beta = 10.0$, $\lambda = 10.0$. I completed one run of 10,000,000 MCMC steps for both groups of 5 individuals and assessed convergence using TRACER (Rambaut et al., 2013). The effective sample size was >200 for posterior, likelihood, prior, and most theta statistics. I discarded the first 10% of trees using TreeAnnotator, then visualized the distribution of species tree topologies using DensiTree (Bouckaert, 2010).

The second analysis, SVDQuartets (Chifman & Kubatko, 2015), is implemented in PAUP (Swofford, 2003) and runs substantially faster than SNAPP. For this reason, I ran the analysis using 10-30 individuals per OTU ($N = 226$). As with SNAPP, I used unlinked, bi-allelic loci and constrained the monophyly of outgroup taxa. All possible quartets were sampled and bootstrapped ($N = 100$).

The MSC analyses cannot handle OTUs with missing data, so using individuals as taxonomic units excludes most SNPs in the dataset. I assigned individuals *a priori* to populations to circumvent this issue. To assess the genetic relationship among individuals without pre-defined population assignments, I constructed a neighbor-joining (NJ) tree with a Kimura 2 substitution model using the program BioNJ (available at *phylogeny.fr*; Dereeper et al., 2008).

3.3.5 Coloration Data from Reflectance Spectroscopy

The color data, derived from reflectance spectroscopy, is described in detail in chapter 2. Briefly, I collected three measurements from the dorsum of five *O. pumilio* individuals per collecting site. I calculated individual-level colorimetrics (total brightness and chroma) from the spectra then performed model-based clustering. Finally, I assigned individuals to color groups based on their cluster membership.

3.3.6 Tests of Phylogenetic Signal and Ancestral Character State Reconstructions of Skin

Coloration

To analyze how lineages resemble each other in coloration, I followed the *phylo.signal.disc* procedure for estimating the phylogenetic signal of discrete data (Paleo-

López et al., 2016). The algorithm estimates the minimum number of character-state transitions at each node that account for the observed distribution of the character in the phylogeny, assuming maximum parsimony (Maddison & Maddison, 2000). Then it compares the observed distribution with the median of a randomized distribution ($N = 1000$). A significant phylogenetic signal is inferred when the observed number of transitions are significantly less than the randomized median.

My reflectance spectroscopy dataset does not include the 'PODO' population from Isla Popa or the ancestral red population from Puerto Viejo, so these taxa were pruned from the species trees for both dorsal and ventral tests of phylogenetic signal. I repeated the *phylo.signal.disc* procedure for the two SNAPP and one SVDquartets topologies.

I performed ancestral character state reconstructions using a polymorphic evolution model (*fitpolyMk*) implemented in the *phytools* package for *R* (Revell, 2012). To account for the presence of multiple dorsal colors in the same population, terminal taxa are allowed multiple character states (e.g. a red/yellow polymorphic population is assigned B+C). For each topology, I fit symmetric, equal, and 'all rates different' (ARD) transition models with unordered character states to my dorsal coloration data, with the root fixed to the 'red' character state. I plotted the probability of character states at internal nodes and visually inspected each model's performance.

3.4 Results

3.4.1 Population Differentiation

Visual inspection of missingness does not suggest systematic differences among collecting sites (Appendix Figure 4). The coancestry matrix from the fineRADstructure analysis shows clear separation between most collecting sites from different islands and regions (Fig. 15). However, several collecting sites from the same island (Isla Colón: BDD, BLB, LAG) or region (La Gloria: GLB, SIL; Aguacate Peninsula: DBP, RA, PMF) show little to no structure, and therefore they are not supported as distinct populations. These collecting sites are grouped in subsequent analyses, such that the 22 *O. pumilio* collecting sites are represented by 16 populations.

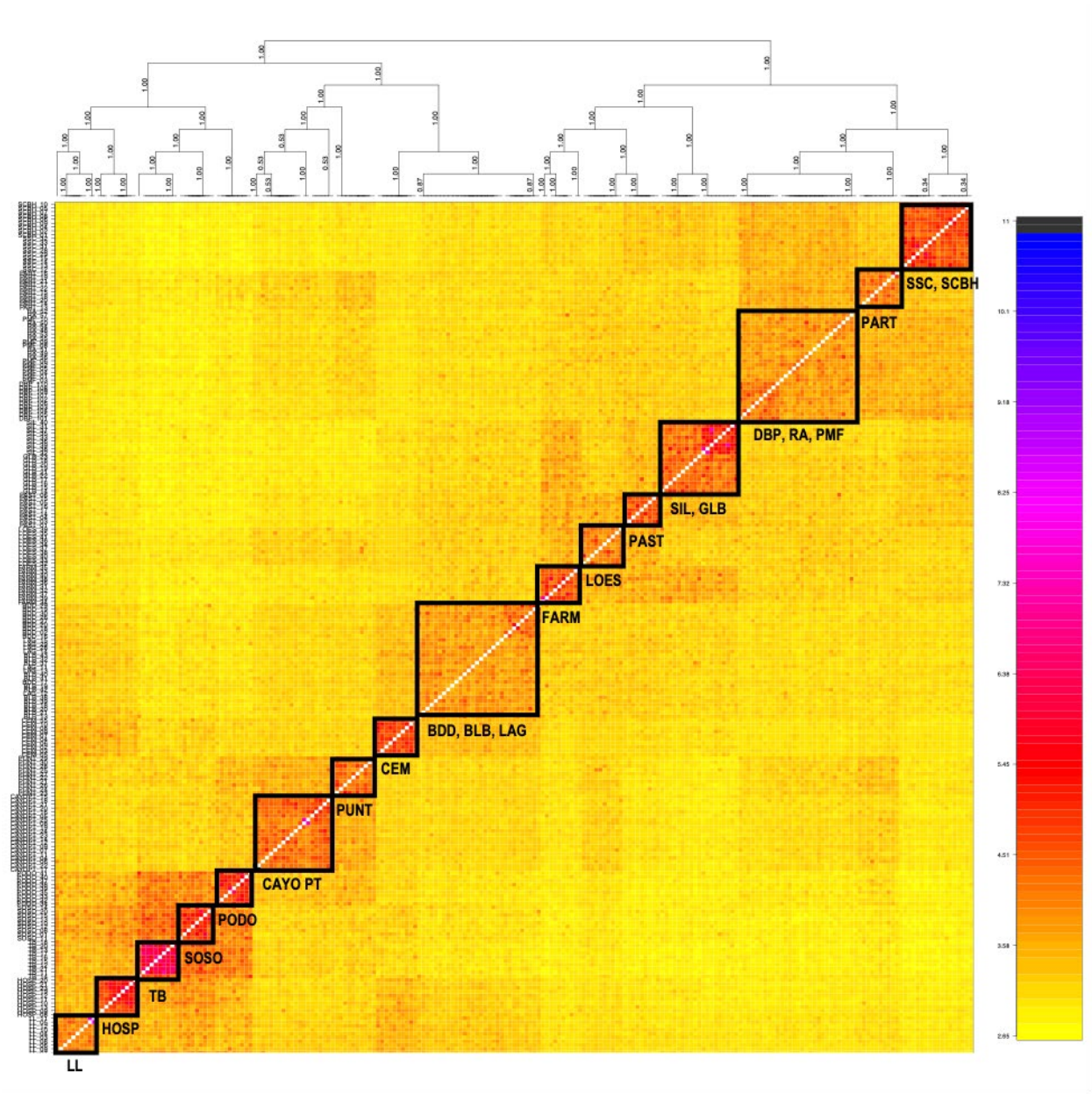


Figure 15. Coancestry matrix from fineRADstructure based on 673 haplotypes containing 2,489 SNPs from 222 *O. pumilio* individuals.

Pairwise coefficients of coancestry are color coded from low (yellow) to high (blue). Black squares outline the 16 OTUs identified in this analysis and are labeled with the collecting sites comprising an OTU. Four OTUs are made up of individuals from >1 collecting site.

3.4.2 Phylogenetic Analysis

The NJ tree, constructed using SNP data from 235 individuals without *a priori* population assignment, supports monophyly of most populations (Appendix Figure 5). Together with the coancestry matrix, this result affirms the *a priori* population assignments provided for subsequent analyses. Furthermore, the genetic relationships estimated by the NJ tree support many relationships estimated by MSC approaches.

For the sake of simplicity, I report the results from one SNAPP run (one set of 5 individuals; Fig. 16, Appendix Figure 6) and the SVDQuartets analysis (Appendix Figure 7). Despite running the SNAPP analysis to 10,000,000 MCMC and bootstrapping the SVDQuartets with 100 replicates, most relationships have low support. However, the topologies resulting from these MSC analyses are largely consistent. Both topologies show the same divergence pattern among CAYO PT, PUNT, HOSP, PODO, SOSO, TB, and LL. Likewise, the sister relationship between COLON and CEM is supported by both topologies. AGUACATE + CRISTOBAL + PART form a group in both species trees, although their groups support different sister taxa. Four populations on or nearest the mainland (GLBSIL, FARM, LOES, and PAST) present major differences between analyses. In particular, PAST and LOES show divergence patterns that have implications for ancestral state reconstructions (below).

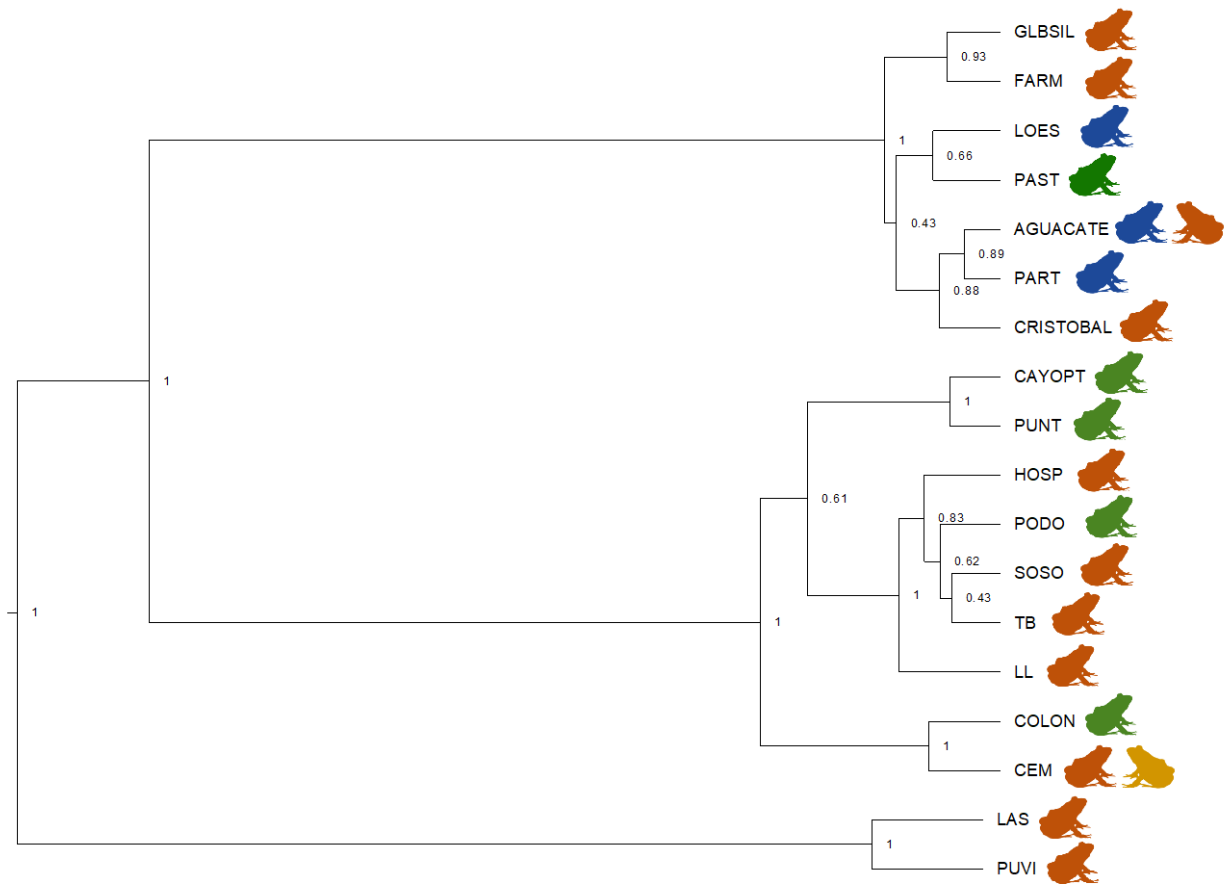


Figure 16. Species tree resulting from a SNAPP analysis that included unlinked, bi-allelic SNPs from 5 individuals per ingroup OTU.

The analysis ran for 10,000,000 MCMC steps and 10% of trees were subsequently discarded as burn-in. Internal nodes are labeled with posterior probabilities. Frog silhouette colors correspond to population-level by-eye color assignments; polymorphic populations (CEM, AGUACATE) are indicated by >1 silhouette showing the extremes of color variation.

3.4.3 Tests of Phylogenetic Signal and Ancestral Character State Reconstructions

The phylogenetic signal of dorsal or ventral coloration was not significant based on the *phylo.signal.disc* procedure ($p_{D/SNAPP1} = 0.053$, $p_{D/SNAPP2} = 0.40$, $p_{D/SVDQ} = 0.415$, $p_{V/SNAPP1} = 0.647$, $p_{V/SNAPP2} = 0.672$, $p_{V/SVDQ} = 0.209$).

The ancestral state reconstructions that incorporated the ARD model resulted in internal nodes with the highest probability for a single character. Of the three topologies, the SVDquartets phylogeny estimated the highest probability for a single character state at internal nodes (Appendix Figure 8). In this ancestral state reconstruction, few nodes were estimated at a probability > 50% for a single character state, so the results should be interpreted with caution. With this in mind, the analysis suggests the convergent evolution of green and blue dorsal coloration, with a single reversal to bright, red coloration.

The SNAPP topologies produced internal nodes with more uncertainty than the SVDquartets phylogeny (topology 1 shown, Fig. 17). Although it is impossible to conclude the evolutionary history of many lineages from this reconstruction, it suggests less convergence on dull coloration, but more reversals to bright colors, than the SVDQuartets topology.

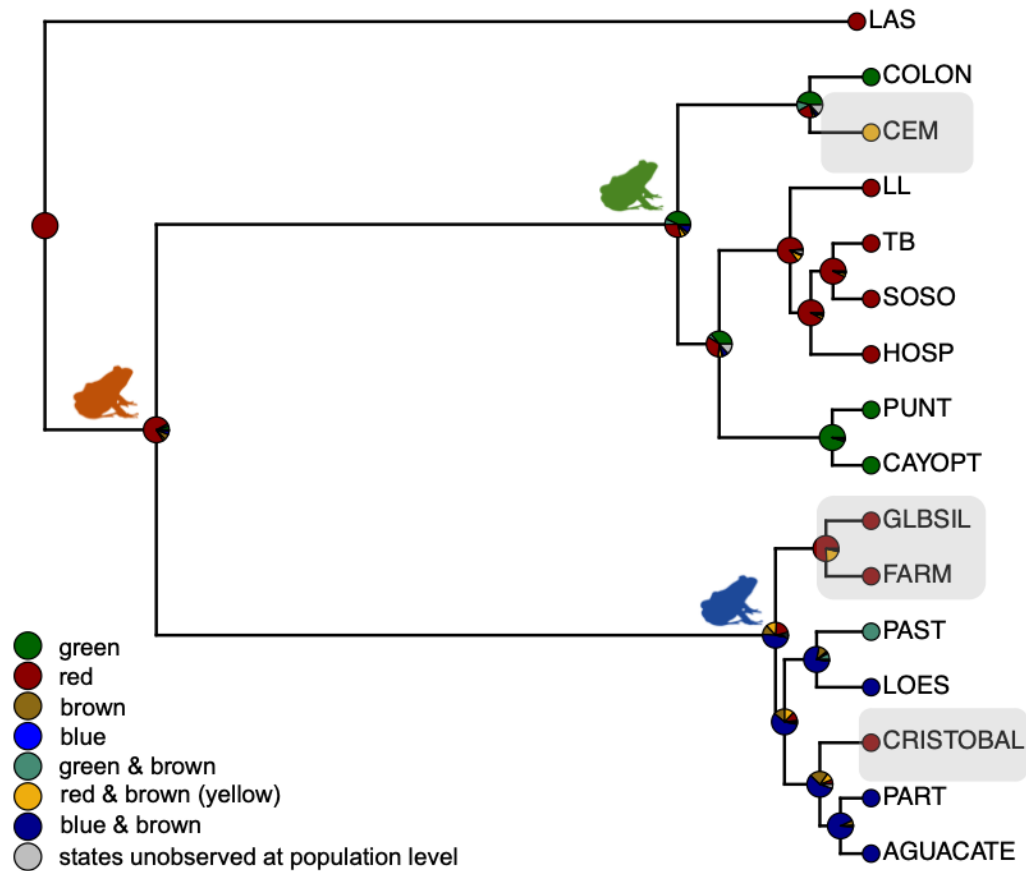


Figure 17. Ancestral character state reconstructions of dorsal coloration using a polymorphic evolution with ‘all rates different’ model and the SNAPP topology.

To account for the presence of multiple dorsal colors in the same population, terminal taxa are allowed multiple character states. Frog silhouettes indicate the character state with the highest probability at internal nodes.

The character reconstruction suggests convergence of dull (green and blue) coloration, with multiple reversals to bright coloration (shaded nodes).

3.5 Discussion

A question of substantial interest in evolutionary biology has been which evolutionary processes are responsible for generating phenotypic divergence. The widespread application of

next-generation sequencing (NGS) to phylogenetic and phylogeographic studies of non-model organisms recently allowed evolutionary biologists to ask this question in systems where the inference had been previously limited by recent timescales of divergence and incompletely sorted lineages. This represents a significant advancement in the field of evolutionary biology because incipient speciation, a process rarely observable empirically, is often confounded by rapid phenotypic divergence and incomplete lineage sorting. Studies of *O. pumilio* provide an excellent opportunity to study the historical, genomic, and physiological underpinnings of trait divergence and potentially incipient speciation. While this species is the focus of many theoretical and empirical studies that assume rapid phenotypic divergence, previous attempts at understanding the evolution of warning coloration in this species have relied on phylogeographic trees with sparse sampling and very limited resolution.

To bridge this gap, I implemented a MSC approach to the analysis of a RADseq dataset obtained from dense sampling of individuals and populations across the Bocas del Toro archipelago. The SNAPP and SVDQuartets topologies derived from extensive sampling across the archipelago and RADseq provide much better resolution than previous studies. Both species trees show low support values, but the relationships among populations are largely consistent.

The SNAPP species tree consists of two main groups that generally correspond to island or mainland populations. The 'island' clade is identical in its relationships to the SVDQuartets tree, except for the latter's inclusion of LOES. The 'mainland' clade includes all the inland populations, the Aguacate Peninsula, and nearby Isla Pastores (PAST). Blue and blue/green morphs are restricted to this part of the tree, leading to different conclusions regarding the evolution of dorsal coloration when compared to the SVDQuartets topology (Figure 17, Appendix Figures 8; discussion below). However, an AGUACATE + CRISTOBAL + PART group is present in both

MSC topologies, though their sister taxa differ. There are three points of consideration for this group. First, the coancestry matrix supports its separation into three populations, but the NJ tree (Appendix Figure 5) suggests less clear relationships. In the dendrogram, AGUACATE is paraphyletic with respect to SAN CRISTOBAL, indicating that some individuals from RA and DBP are more closely related to San Cristobal populations. Second, the separation of San Cristobal and Isla Partida from the Aguacate Peninsula is ~1000 years old, so resolving genetic relationships in this region will be especially challenging. Finally, the exact arrangement of taxa within this group leads to different conclusions about the evolution of red/blue polymorphism in *O. pumilio*.

Interpreting the species trees requires caution, as support values are low for most lineages. Nevertheless, the distribution of dorsal coloration across the phylogeny (Fig. 16, Appendix Figure 7) hints at the relative roles of selective versus neutral processes in shaping genetic diversity in *O. pumilio*. If selection on dorsal coloration overwhelmed signals of the region's geological history, this would be reflected in the species tree by clustering of morphs. Namely, green frogs from the geographically distant Isla Colon and Isla Cayo Agua would be more closely related to each other than to populations from nearby islands. Likewise, red frogs from Isla Bastimentos and Isla Solarte would be most closely related to mainland red frogs. In this study, *O. pumilio* color morphs are not co-localized on the phylogeny (Fig. 16, Appendix Figure 7), suggesting that the region's geological history, not selection on dorsal coloration, has shaped genetic diversity across the islands.

A vicariance model for *O. pumilio* evolution can be probed using the historical reconstructions of island isolation (Anderson & Hadley, 2002; Wang & Shaffer, 2008). Since the topological relationships predicted by a vicariance model change from the start (~8,900 years ago) to the end (~1,000 years ago) of island formation (Wang & Shaffer, 2008), we gain insight into *O.*

pumilio evolution based on the spatial patterns of lineages alone. For example, if the configuration of islands at 5,200 years ago explains the relationships among lineages, I would expect populations connected at those times to be most closely related (Fig. 18A). In contrast, if the current island configuration explains the relationships among lineages, I expect populations from the same island to be most closely related (Fig. 18B). Examination of the 'island' clade supports the former scenario (Fig. 18C). PODO, one of two populations I sampled on Isla Popa, is sister to a group containing populations from Isla Solarte (SOSO) and Isla Bastimentos (TB). The second of two populations located on Isla Popa, PUNT, is sister to Isla Cayo Agua (CAYO PT), reflecting those populations' connectivity until 3400 years ago. Similarly, the polymorphic Isla Bastimentos population (CEM) is sister to Isla Colón, reflecting these populations' connectivity from before the islands' formation; another Isla Bastimentos population (LL) is sister to a group containing Isla Popa + Isla Solarte + Bastimentos, populations connected up to 5200 years ago.

Phylogenetic signal is the propensity of related lineages to resemble each other more than a random distribution of lineages from the same phylogeny (Münkemüller et al., 2012). In this study, tests of phylogenetic signal for dorsal and ventral coloration data were not significant. This result is not surprising given the prevalence of convergent evolution indicated by a previous analysis (Wang & Shaffer, 2008). The ancestral character state reconstructions in this study further support the occurrence of convergent evolution of coloration. Although there is high uncertainty at internal nodes for SVDQuartets and SNAPP topologies, both reconstructions agree on the convergent evolution of dull colors like green and blue and at least one reversal to bright coloration like yellow, orange or red. Future studies of coloration in this species should explore the selective pressures driving parallel evolution and the molecular mechanism of convergence, particularly involving polymorphic populations.

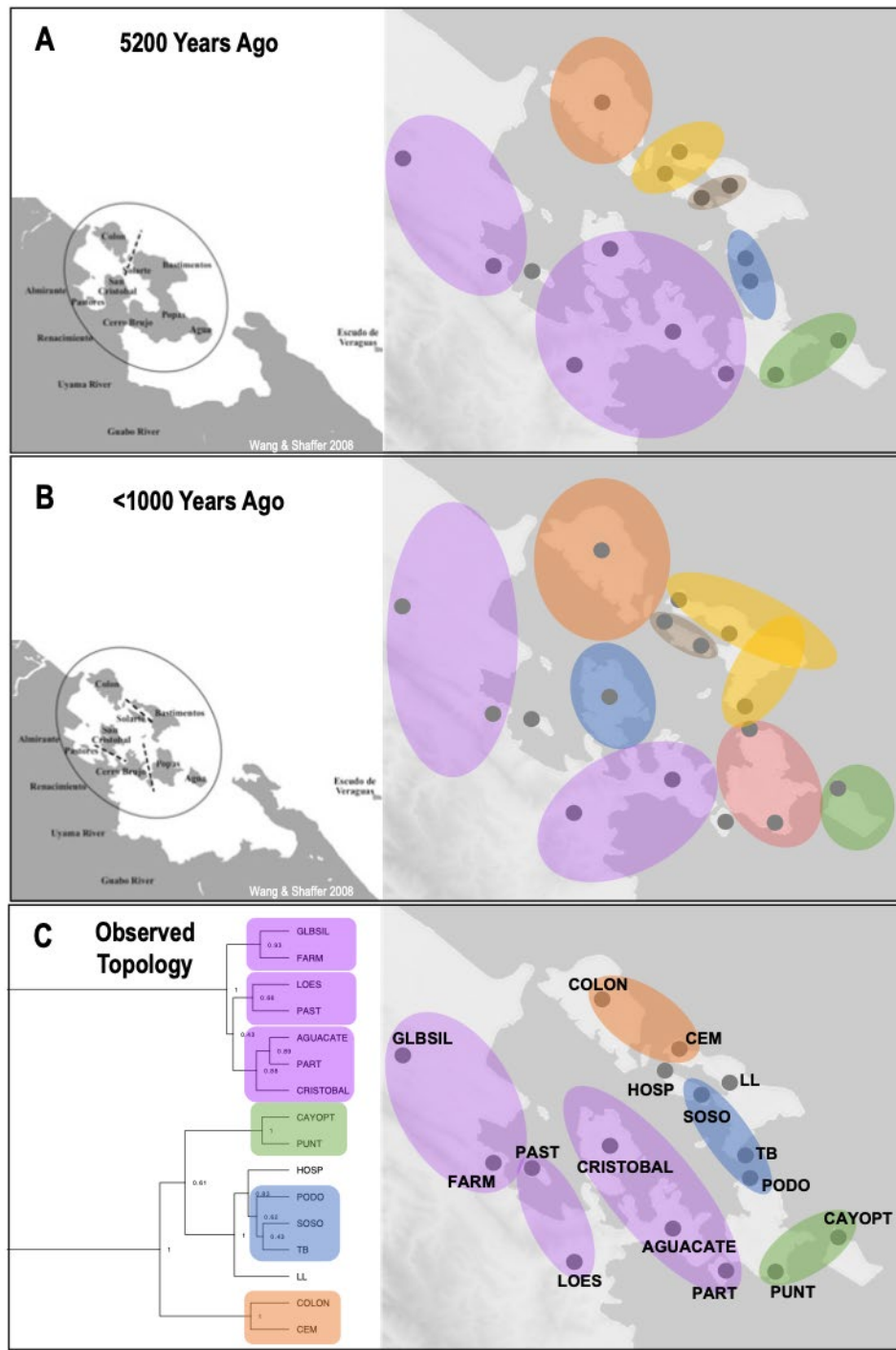


Figure 18. Predicted topologies based on vicariance models (A, B) and the observed topology (C).
The topological relationships predicted at 5,200 years ago (A) differs from the topological relationships predicted by the current configuration (B). The former model expects relationships to reflect previous

connectivity (yellow, brown, blue, and green areas in (A)), while the latter model expects populations on the same island to be more closely related than to populations on different islands (red, yellow, and brown areas in (B)). The observed topology (C), supports the model of vicariance predicted at 5,200 years ago, as the relationships reflect the connectivity at that time.

Oophaga pumilio has emerged as a popular species for evolutionary biologists interested in the roles of geographic isolation and selection in driving phenotypic divergence. However, researchers have been unable to resolve the evolutionary processes responsible for warning color variation in this species. Previous molecular analyses of this system sampled sparsely throughout the archipelago and were based on markers often ill-suited for investigations of recent population divergences. Using RAD tags, I generated thousands of genome-wide markers in *O. pumilio*, far beyond what has been achieved with previous molecular techniques. This approach gave me an unparalleled ability to resolve genetic differences and estimate phylogeographic relationships among *O. pumilio* populations. The phylogenies resulting from this analysis will provide a backbone for ongoing genetic, chemical, behavioral, and ecological studies of phenotypic evolution in this species, which is a promising system for the study of phenotypic evolution.

Appendix A Supplement to Chapter 1

Appendix Table 1. Marker segregation types used in the construction of linkage maps and the mapping of candidate loci.

Each nucleotide is color-coded to illustrate the information provided by the marker. Male-informative markers are heterozygous in the male parent, female-informative markers are heterozygous in the female parent, partially informative markers are heterozygous in both parents but possess the same combination of nucleotides. Fully informative markers are heterozygous in both parents and possess different combinations of nucleotides, allowing for the full phasing of loci.

Segregation Type	Male Parent Genotype	Female Parent Genotype	Information Provided
lm x ll	heterozygous AG	homozygous GG	male informative
nn x np	homozygous GG	heterozygous AG	female informative
hk x hk	heterozygous AG	heterozygous AG	partially informative
ef x eg	heterozygous AG	heterozygous AC	fully informative
ab x cd	heterozygous AG	heterozygous TC	fully informative

Appendix Table 2. Forward and reverse primer sequences, annealing temperatures, and the approximate amplified fragment length for five candidate loci.

Primer Pair Name	Forward Primer	Reverse Primer	Annealing Temperature (T _A), °C	Approximate Amplified Fragment Length (bp)
c04_O	5'-CCAAGGACTGGGT TTTCCTAAACTA-3'	5'-ATCTCTGTAATCCC GTCATGGAAC-3'	53	500
c10_E	5'-GTCGGCTCCCTGT GTATATG-3'	5'-ATCTCTCGGCTTAC TGGGGT-3'	59	980
c12_E	5'-CAAGAACTGTTTG GCCCCG-3'	5'-CTCACAGCTCACA CAACCCT-3'	59	600
c15_H	5'-GGCGAACAGGATC ACAAGGA-3'	5'-GCGTTTCGTGTGA GCTTCTT-3'	59	830
c16_E	5'-CTGGGAAGGGGGA GGGATAA-3'	5'-ACCGGCAGGCAT GGTAATAG-3'	53	820

Appendix Table 3. Candidate regions for markers heterozygous in one or more parents.

The segregation type and the number of offspring contributing SNPs to each marker are also listed.

Candidate Region	Informative Marker ID	Segregation Type	Number of Offspring Contributing
c16_E	IM57	lm x ll	31
c12_E	IM43	ef x eg	11
c12_E	IM41	lm x ll	54
c12_E	IM46	hk x hk	13
c12_E	IM48	nn x np	15
c10_E	IM58	nn x np	41
c10_E	IM61	lm x ll	15
c10_E	IM63	hk x hk	10
c10_E	IM67	lm x ll	16
c15_H	IM33	lm x ll	14
c15_H	IM22	hk x hk	44
c04_O	IM03	hk x hk	14
c04_O	IM05	hk x hk	15
c04_O	IM16	nn x np	44
c04_O	IM15	nn x np	43
c04_O	IM09	nn x np	28
c04_O	IM08	lm x ll	44
c04_O	IM07	lm x ll	41

Appendix B Supplement to Chapter 2

Appendix Table 4. Sampling locations, abbreviations, and GPS coordinates included in the histological and HPLC datasets.

Abbreviation	Sampling Site	Longitude	Latitude
LL	La Loma, Isla Bastimentos	9°19'1.60"N	82° 9'0.19"W
HO	Hospital Point, Isla Solarte	9°19'55.50"N	82°13'5.67"W
TB	Tranquilo Bay, Isla Bastimentos	9°15'16.90"N	82° 8'47.90"W
SS	southeast Isla Solarte	9°18'27.01"N	82°10'42.27"W
NP	north Isla Popa	9°13'40.99"N	82° 7'11.91"W
CA	Cayo Point, Isla Cayo Agua	9°10'48.60"N	82° 3'4.20"W
PU	Punta Laurel, Isla Popa	9°8'28.90"N	82° 7'34.37"W
CE, CY	Cemetery Hill, Isla Bastimentos	9°20'56.58"N	82°12'34.47"W
OL	Old Point, Isla Bastimentos	9°17'53.50"N	82° 5'10.70"W
IC	Isla Colón	9°24'29.94"N	82°17'58.05"W
FA	Farm, mainland	9°13'27.60"N	82°22'3.90"W
LO	Loma Estrella, mainland	9°10'17.60"N	82°18'28.30"W
NI	Isla Pastores	9°14'24.76"N	82°20'37.46"W
SI	El Silencio, mainland	9°23'32.40"N	82°32'4.60"W
GL	La Gloria Bajo, mainland	9°10'17.60"N	82°18'28.30"W
DB, DI, DR	Dolphin Bay Preserve, Aguacate peninsula	9°13'14.77"N	82°13'3.01"W
NR	Rana Azul, Aguacate peninsula	9°10'42.10"N	82°16'0.80"W
PM	Loma Partida, Aguacate peninsula	9° 9'1.40"N	82°11'4.70"W
SC	Isla San Cristóbal	9° 16'21.8"N	82°15'6.9"W
PA	Isla Loma Partida	9° 8'57.30"N	82°10'5.00"W
CR	La Selva Biological Station, Costa Rica	10° 25'47.0"N	84°00'10.8"W
SSC	south Isla San Cristóbal	9°13'40.70"N	82°16'28.20"W
PODO	Popa Dos, Isla Popa	9°13'4.90"N	82° 8'21.60"W
BLBLAG	east Isla Colón	9°23'42.60"N	82°16'32.70"W

Appendix Table 5. Results of linear mixed effects models with total pigment as the predictor variable and color metric as the explanatory variables.

Explanatory Variable (x)	Fixed Effects Slope Value	Standard Error	p-value	AIC	Marginal R ²	Conditional R ²
B_I	-0.0476792	0.03216465	0.142	-321.73	0.02252569	0.9378994
S1_V	0.1410919	0.08716797	0.1093	-324.1358	0.02551675	0.9378129
S1_B	0.0163282	0.02223718	0.4648	-319.3388	0.005354288	0.9372769
S1_G	0.0173216	0.06806765	0.7997	-321.0988	0.0006728555	0.9370875
S1_Y	-0.2356148	0.15681623	0.1367	-324.9589	0.02245263	0.9379553

S1_R	-0.0636407	0.04992442	0.2059	-322.0392	0.01592084	0.9374584
-----------------------	------------	------------	--------	-----------	------------	-----------

Appendix Table 6. Summary of similarity percentages breakdown (SIMPER) test.

Asterisks indicate the level of significance for pairwise contrasts. An empty cell implies non-significance.

Pairwise color contrasts without any significant differences are not shown.

Carotenoid Pigment	Pairwise Color Group Contrast			
	blue - red	blue - green	blue - brown/'other'	red - green
Apocarotenoid		**		
Xanthophyll				
Cis Xanthophyll				
Canary Xanthophyll				
Canary Xanthophyll Ester 1		***		
Canary Xanthophyll Ester 2		***		
Canary Xanthophyll Ester 3		*		
Echinenone			***	
X3 Hydroxy Echinenone				
Canthaxanthin				
Canthaxanthin Ester 1	*		***	
Canthaxanthin Ester 2			*	
β-Carotene	*			*
Ketocarotenoid Ester 1				
Ketocarotenoid Ester 2		**		
Cis Lutein Ester	*			*
Lutein Ester 1				
Unknown 1				
Unknown 2				

Appendix Table 7. Detailed summary of linear models of the relationship between individual carotenoid pigments and color metric variable.

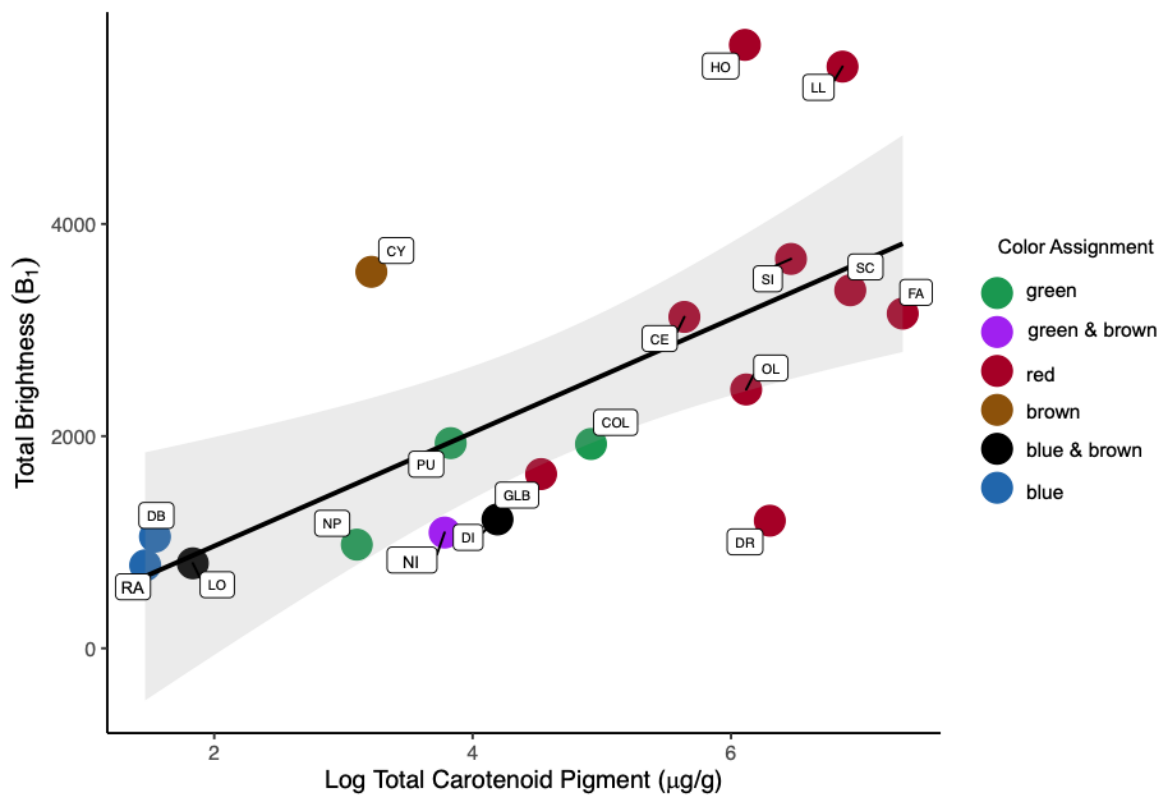
Entries list slope, t value, p values, and R²_{adj}.

Significance codes: 0 '***' 0.001 '**' 0.01 '*' not significant 'ns'

Carotenoid Pigment	Model Stat.	B ₁	S1 _R	S1 _Y	S1 _G	S1 _B	S1 _V
Apocarotenoid	slope		0.04941		- 0.031749		
	std. error		0.01516		0.006808		
	t value		3.258		-4.664		
	p value		0.004934 **		0.000259 ***		
	R ² _{adj}		0.3613		0.5497		
Canary Xanthophyll	slope	14.940	0.0018474			-0.0008016	- 0.0006030

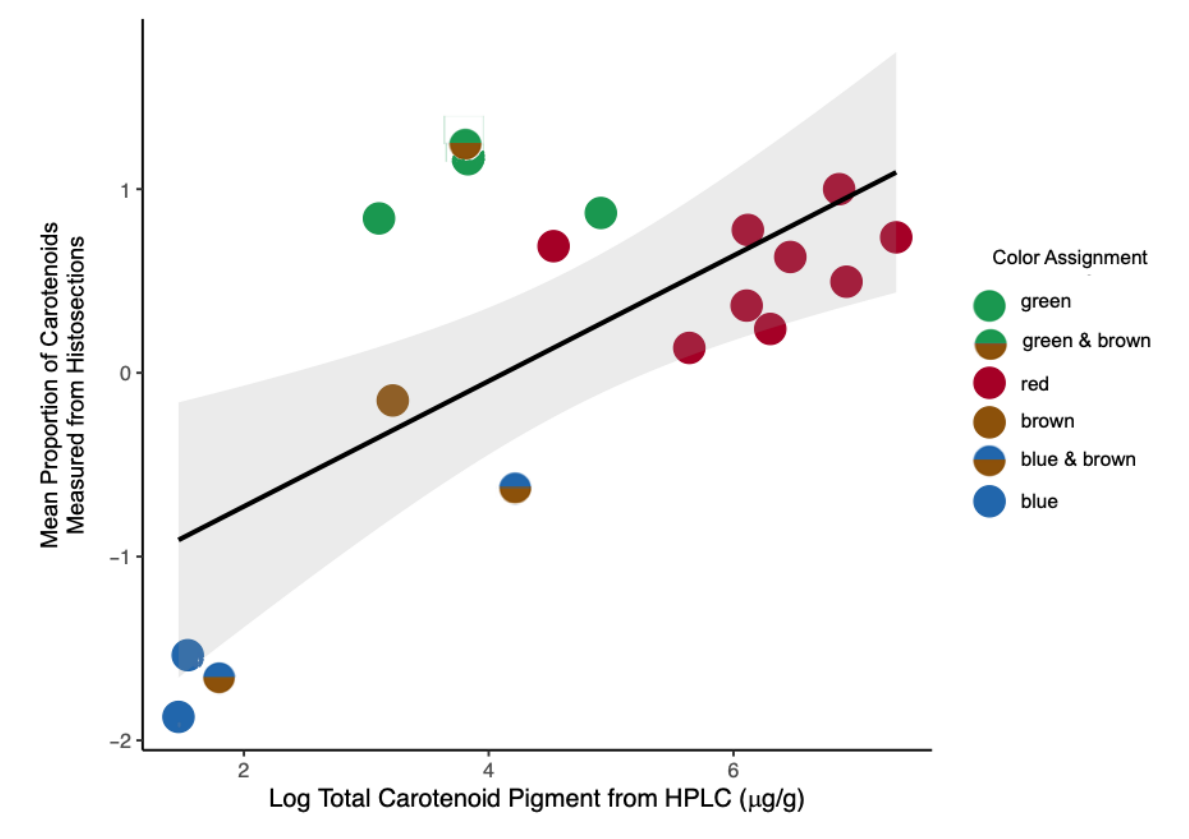
Ester 1	std. error	6.686	0.0006234			0.0003596	0.0002436
	t value	2.235	2.964			-2.229	-2.475
	p value	0.040066 *	0.00915 **			0.0405 *	0.0249 *
	R²_{adj}	0.1902	0.314			0.1893	0.2316
Canary Xanthophyll Ester 2	slope	13.948	0.0016058			-0.0007182	-
	std. error	5.081	0.0004781			0.0002784	0.0001872
	t value	2.745	3.358			-2.580	-2.874
	p value	0.014378 *	0.004 **			0.0201 *	0.011 *
	R²_{adj}	0.2777	0.3768			0.2497	0.2992
Canary Xanthophyll Ester 3	slope		0.004227			-0.0017550	-
	std. error		0.001382			0.0008121	0.0005524
	t value		3.059			-2.161	-2.375
	p value		0.0075 **			0.0462 *	0.0304 *
	R²_{adj}		0.3296			0.1776	0.2144
Echinenone	-						
Canthaxanthin Ester 1	slope		0.0021189				
	std. error		0.0007741				
	t value		2.737				
	p value		0.0146 *				
	R²_{adj}		0.2764				
Canthaxanthin Ester 2	slope		0.003488				
	std. error		0.001323				
	t value		2.637				
	p value		0.0179 *				
	R²_{adj}		0.2593				
β-Carotene	slope	8.859	0.0009177			-0.0003952	-3.096e-04
	std. error	2.003	0.0001972			0.0001265	8.057e-05
	t value	4.422	4.654			-3.124	-3.842
	p value	0.000427 ***	0.000265 ***			0.00654 **	0.00144 **
	R²_{adj}	0.5219	0.5485			0.3401	0.4474
Ketocarotenoid Ester2	slope	28.93	0.003289	0.0010737		-0.0015563	-
						0.0011713	

	std. error	12.59	0.001220	0.0004504		0.0006767	0.0004576
	t value	2.298	2.697	2.384		-2.300	-2.560
	p value	0.035396 *	0.0159 *	0.0299 *		0.0352 *	0.021 *
	R²_{adj}	0.2011	0.2695	0.2159		0.2015	0.2462
Cis Lutein Ester	slope	83.48	0.008939			-0.003653	- 0.0028912
	std. error	18.16	0.001691			0.001173	0.0007406
	t value	4.598	5.286			-3.113	-3.904
	p value	0.000297 ***	7.38e-05 ***			0.0067 **	0.00126 **
	R²_{adj}	0.5423	0.6132			0.3383	0.4558
Lutein Ester 1	slope	40.44	0.004143			-0.0018532	- 0.0014406
	std. error	13.64	0.001374			0.0007864	0.0005233
	t value	2.964	3.016			-2.356	-2.753
	p value	0.3141**	0.00819 **			0.0315 *	0.0142 *
	R²_{adj}	0.00915	0.3227			0.2112	0.279



Appendix Figure 1. Scatterplot illustrating the relationship between total brightness (B_1) and total carotenoids measured from HPLC.

The linear regression indicate total carotenoids from HPLC explain 41.3% of the variance ($R^2_{\text{adj}} = .413$, $F(1,16) = 12.97$, $p = .0024$) and predict total brightness ($t = 6.02$, $p = 0.00239$).



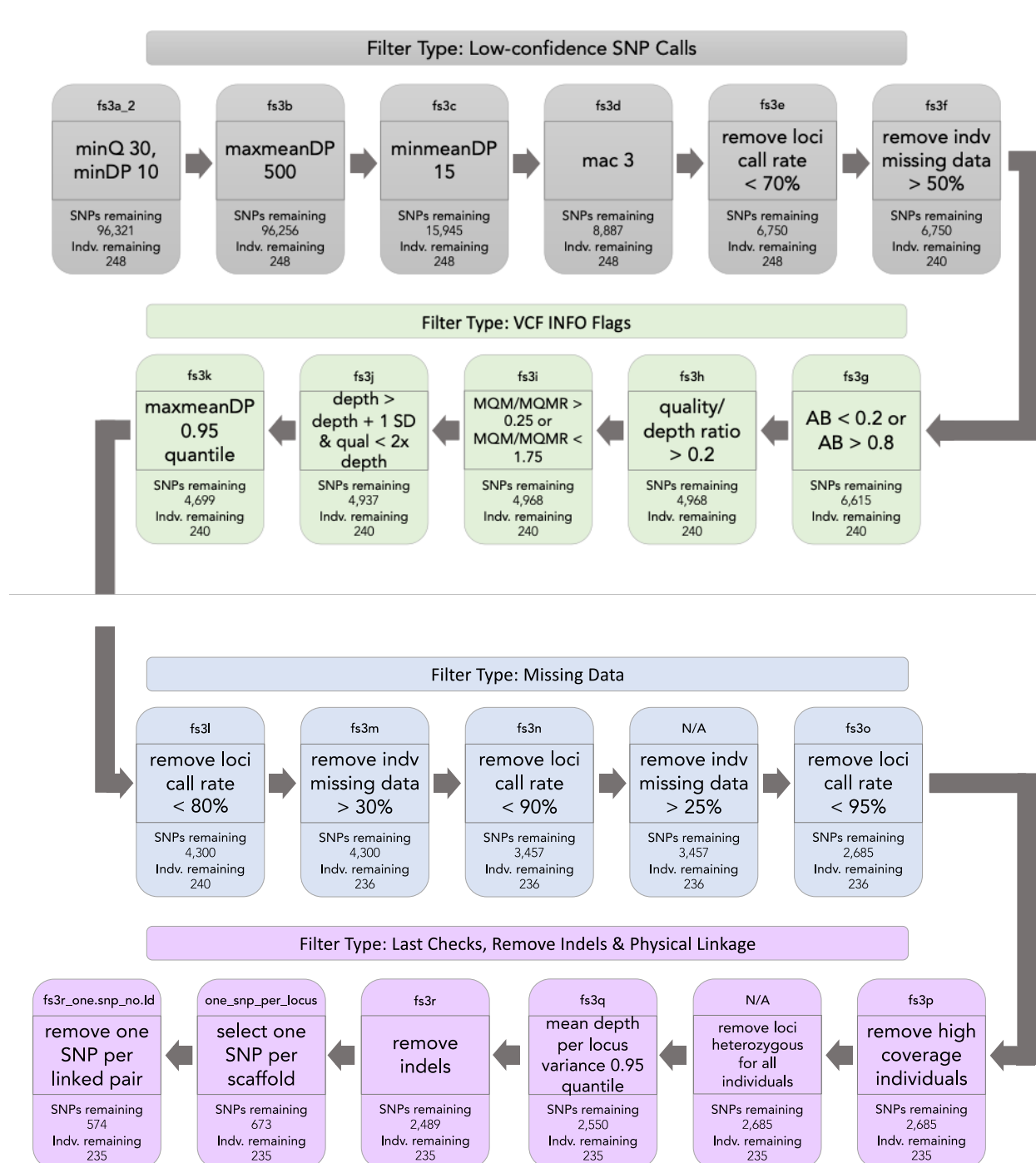
Appendix Figure 2. Scatterplot illustrating the relationship between the proportion of carotenoids measured from histosections and the total carotenoids measured from HPLC.

The linear regression indicates that total carotenoids from HPLC predict carotenoid proportion measured from histological sections ($t = 3.59, p = 0.00247$) and explains 41.1% of its variation ($R^2_{adj} = 0.411, F(1,16) = 12.86, p = 0.0025$).

Appendix C Supplement to Chapter 3

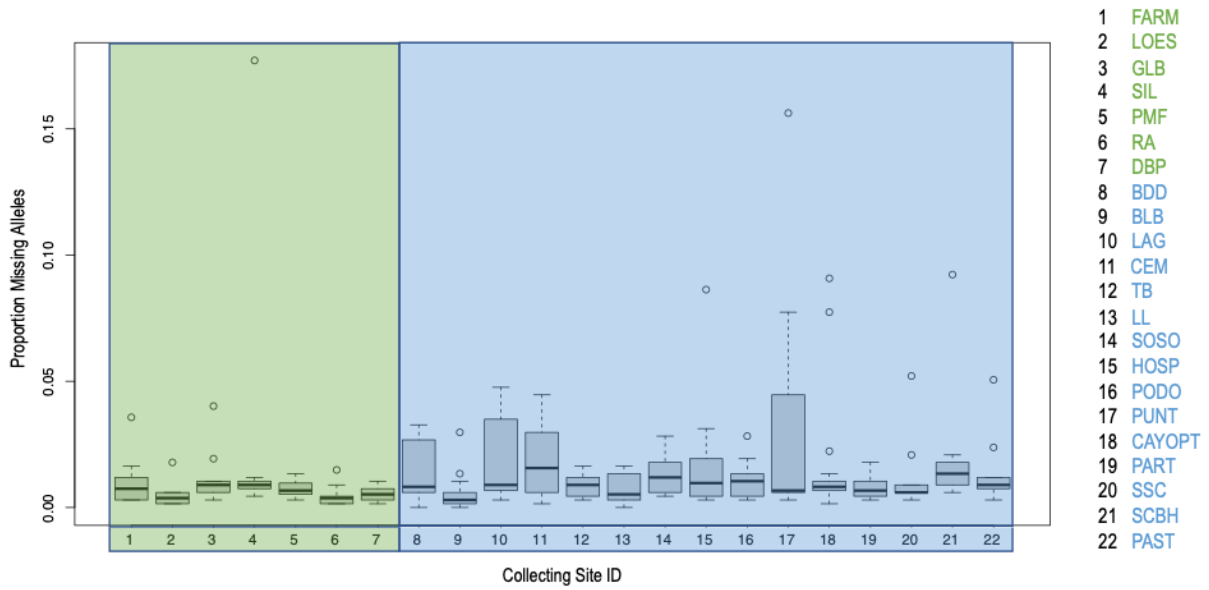
Appendix Table 8. Sampling locations, abbreviations, and GPS coordinates included in the phylogenetics dataset.

Abbreviation	Sampling Site	Longitude	Latitude
LL	La Loma, Isla Bastimentos	9°19'01.6"N	82°09'19.1"W
HOSP	Hospital Point, Isla Solarte	9°19'55.5"N	82°13'08.4"W
TB	Tranquilo Bay, Isla Bastimentos	9°25'43.2"N	82°14'56.6"W
SOSO	southeast Isla Solarte	9°18'01.9"N	82°10'18.2"W
PODO	Popa Dos, Isla Popa	9°13'4.90"N	82° 8'21.60"W
CAYOPT	Cayo Point, Isla Cayo Agua	9°10'48.60"N	82° 3'4.20"W
PUNT	Punta Laurel, Isla Popa	9°08'28.9"N	82°07'37.3"W
CEM	Cemetery Hill, Isla Bastimentos	9°20'56.58"N	82°12'34.47"W
BDD	Boca del Drago, Isla Colón	9°25'21.6"N	82°19'15.5"W
BLB	Bluff Beach, Isla Colón	9°24'12.5"N	82°15'05.9"W
LAG	La Gruta, mainland	9°23'42.6"N	82°16'32.7"W
FARM	Farm, mainland	9°13'27.60"N	82°22'3.90"W
LOES	Loma Estrella, mainland	9°10'17.60"N	82°18'28.30"W
PAST	Isla Pastores	9°14'24.76"N	82°20'37.46"W
SIL	El Silencio, mainland	9°23'32.40"N	82°32'4.60"W
GLB	La Gloria Bajo, mainland	9°19'15.8"N	82°29'49.3"W
DBP	Dolphin Bay Preserve, Aguacate peninsula	9°13'14.77"N	82°13'3.01"W
RA	Rana Azul, Aguacate peninsula	9°10'42.1"N	82°16'00.8"W
PMF	Loma Partida, Aguacate peninsula	9° 9'1.40"N	82°11'4.70"W
SSC	south Isla San Cristóbal	9°13'40.7"N	82°16'28.2"W
SCBH	Isla San Cristóbal	9°16'21.96"N	82°15'7.93"W
PART	Isla Loma Partida	9°08'57.3"N	82°10'05.0"W
LAS	La Selva Biological Station, Costa Rica	10° 25'47.0"N	84°00'10.8"W
PUVI	Puerto Viejo de Talamanca, Costa Rica	9°38'51.92"N	82°44'40.16"W



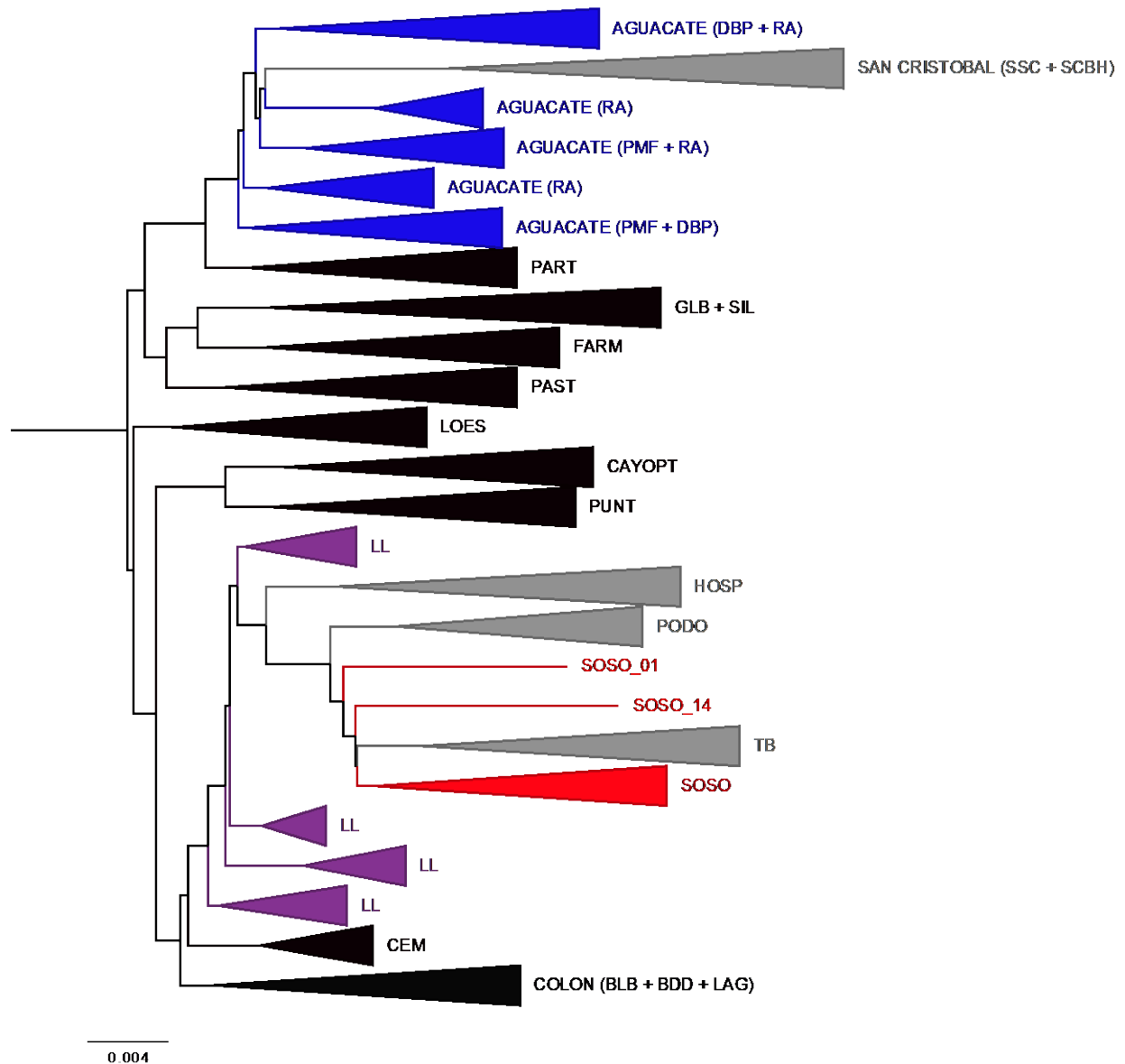
Appendix Figure 3. Schematic detailing the bioinformatic steps taken to filter SNP variants.

The number of SNPs and individuals remaining are listed for each step. Full details are provided in O’Leary et al. (2018).



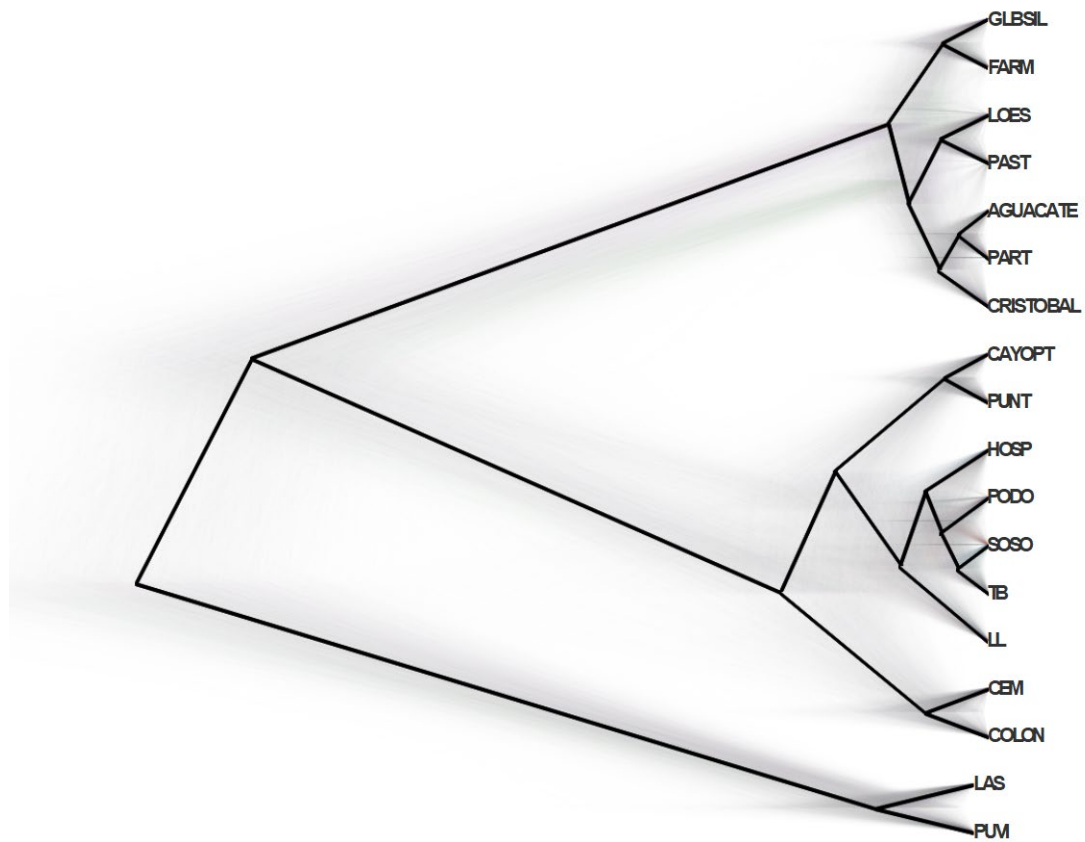
Appendix Figure 4. Boxplot showing the proportion of missing alleles by collecting site.

Mainland collecting sites are shaded green and island collecting sites are shaded blue. Numbers along the x axis correspond to collecting sites, which are listed to the right of the plot. There do not appear to be systematic patterns of missingness between island and mainland collecting sites.



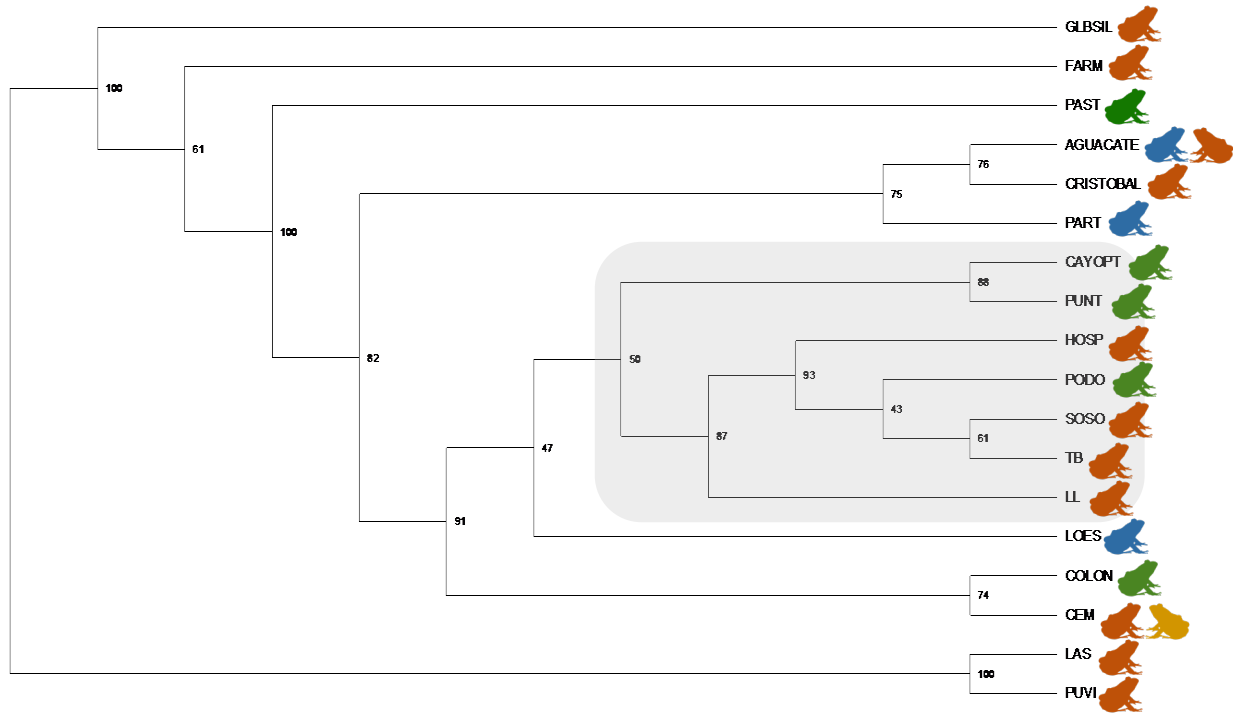
Appendix Figure 5. Neighbor joining tree with a Kimura 2 substitution model, demonstrating reciprocal monophyly among the majority of OTUs identified in the coancestry matrix (Fig. 15).

In this NJ analysis, individuals were not given a priori population assignments. The dendrogram has been rooted using the midpoint. Monophyletic populations are indicated by black and gray nodes. The paraphyletic populations Aguacate, LL, and SOSO are colored blue, purple, and red, respectively.



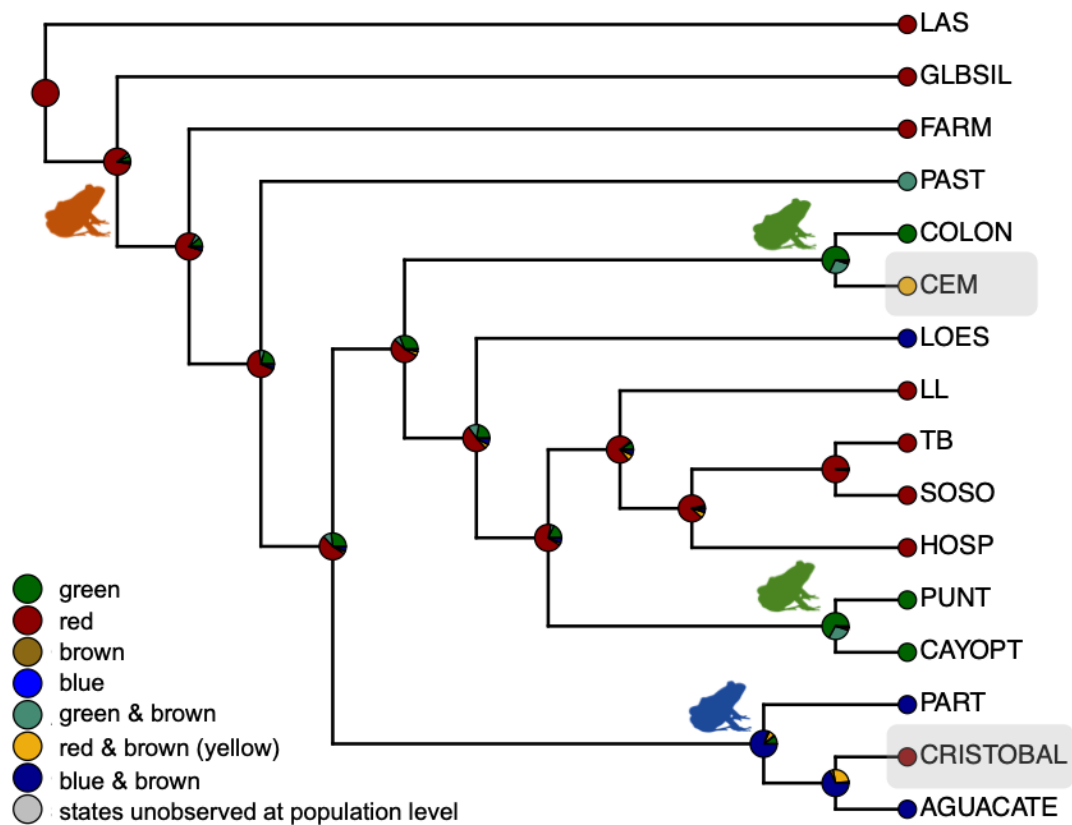
Appendix Figure 6. Densitree resulting from a SNAPP analysis that included unlinked, bi-allelic SNPs from 5 individuals per ingroup OTU.

The analysis ran for 10,000,000 MCMC steps and 10% of trees were subsequently discarded as burn-in. Each light grey line represents one of the MCMC steps. The solid black line indicates the consensus tree.



Appendix Figure 7. Species tree resulting from the SVDQuartets analysis implemented in PAUP and bootstrapped with 100 replicates.

This analysis included 10-30 individuals per OUT (N=226). The shaded lineages indicate the relationships that are identical to the SNAPP topology (Fig. 16).



Appendix Figure 8. Ancestral character state reconstructions of dorsal coloration using a polymorphic evolution with ‘all rates different’ model and the SVDQuartets topology.

To account for the presence of multiple dorsal colors in the same population, terminal taxa are allowed multiple character states. Frog silhouettes indicate the character state with the highest probability at internal nodes.

The character reconstruction suggests convergence of dull (green and blue) coloration, with multiple reversals to bright coloration (shaded nodes).

Bibliography

- Aitchison J. 1982. The statistical analysis of compositional data. *Journal of the Royal Statistical Society, Series B (Statistical Methodology)* 44:139–177.
- Anderson RP, Handley JCO. 2002. Dwarfism in insular slots: biogeography, selection, and evolutionary rate. *Evolution* 56:1045-1058.
- Bagnara JT, Fernandez P, Fujii R. 2007. On the blue coloration of vertebrates. *Pigment Cell Research* 20:14-26.
- Bagnara JT, Taylor JD, Hadley ME. 1968. The dermal chromatophore unit. *Journal of Cell Biology* 38:67-79.
- Bagnara JT, Hadley ME. 1973. *Chromatophores and Color Change*, In Stern HW (ed) *Organismal Biology Series*, Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Bagnara, JT, Ferris WR. 1971. *Interrelationship of chromatophores*, In Kawamura T, Fitzpatrick, TB, and Seiji M. (ed.), *Biology of the normal and abnormal melanocyte*, Univ. of Tokyo Press, Tokyo.
- Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, et al. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376.
- Batista A, Köhler G. 2008. Variation in *Oophaga pumilio* (Amphibia: Anura: Dendrobatidae) in western Panama. *Salamandra* 44:225–234.
- Berns MW, Narayan KS. 1970. A histochemical and ultrastructural analysis of the dermal chromatophores of the variant ranid blue frog. *Journal of Morphology* 132:169-180.
- Bouckaert, R. R. 2010. DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26:1372–1373.
- Box GEP. 1949. A general distribution theory for a class of likelihood criteria. *Biometrika* 36:317–346.
- Bradshaw HDJ, Schemski DW. 2003. Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature* 426:176-178.
- Brelsford A, Toews DP, Irwin DE. 2017. Admixture mapping in a hybrid zone reveals loci associated with avian feather coloration. *Proceedings of the Royal Society, B*. 284: 20171106.

- Briolat ES, Burdfield-Steel ER, Paul SC, Ronka KH, Seymoure BM, Stankowich T, Stuckert AMM. 2019. Diversity in warning coloration: selective paradox or the norm? *Biological Reviews* 94:388-414.
- Britton G. 1995. Structure and properties of carotenoids in relation to function. *FASEB J* 9:1551-1558.
- Brown JL, Maan ME, Cummings ME, Summers K. 2010. Evidence for selection on coloration in a Panamanian poison frog: a coalescent-based approach. *Journal of Biogeography* 37:891-901.
- Bryant D, Bouckaert R, Felsenstein J, Rosenberg NA, Choudhury AR. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. *Molecular Biology and Evolution* 29:1917–1932.
- Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH. 2011. Stacks: building and genotyping loci *de novo* from short-read sequences. *Genes, Genomes, Genetics (G3)* 1:170-182.
- Chan YF, Marks M, Jones FC, Villarreal GJ, Shapiro MD, et al. 2010. Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a *Pitx1* Enhancer. *Science* 5963:302-305.
- Chifman J, Kubatko L. 2014. Quartet inference from SNP data under the coalescent. *Bioinformatics* 30: 3317–3324.
- Clark KR. 1993. Nonparametric multivariate analyses of changes in community structure. *Australian Ecology* 18:117-143.
- Coates AG, McNeill DF, Aubry M-P, Berggren WA, Collins LS. 2005. An introduction to the geology of the Bocas del Toro Archipelago, Panama. *Caribbean Journal of Science* 41(3):374-391.
- Comeault AA, Flaxman SM, Riesch R, Curran E, Soria-Carrasco V, et al. 2015. Selection on a genetic polymorphism counteracts ecological speciation in a stick insect. *Current Biology* 25:1975-1981.
- Crothers L, Saporito RA, Yeager J, Lynch K, Friesen C, Richards-Zawacki CL, et al. 2016. Warning signal properties covary with toxicity but not testosterone or aggregate carotenoids in a poison frog. *Evolutionary Ecology* 4:601-621.
- Daly, JW, Myers CW. 1967. Toxicity of Panamanian poison frogs (*Dendrobates*): Some biological and chemical aspects. *Science* 56(3777):970-973.
- de Birto-Gitirana L, Azevedo RA. 2005. Morphology of *Bufo ictericus* integument (Amphibia, Bufonidae). *Micron* 36:532-538.

- Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acid Research* 36:465-469.
- Dick C, Hinh J, Hayashi CY, Reznick DN. 2018. Convergent evolution of coloration in experimental introductions of the guppy (*Poecilia reticulata*). *Ecology and Evolution* 8:8999-9006.
- Drewes RC, Hillman SS, Putnam RW, Sokol OM. 1977. Water, nitrogen and ion balance in the African tree frog *Chiromantis petersi boulenger* (anura: Rhacophoridae), with comments on the structure of the integument. *Journal of Comparative Physiology Part B* 116:257–267.
- Duellman WE, Trueb L. 1986. *Biology of Amphibians*, McGraw-Hill Book Company, New York, U.S.A.
- Dugas MB, Halbrook SR, Killius AM, del Sol FJ, Richards-Zawacki CL. 2015. Colour and escape behaviour in polymorphic populations of an aposematic poison frog. *Ethology* 121:1-10.
- Egozcue et al. 2003. Isometric log-ratio transformations for compositional data analysis. *Mathematical Geology* 35:279-300.
- Endler JA. 1980. Natural selection on color patterns in *Poecilia reticulata*. *Evolution* 34:76-91.
- Etter PD, Preston JL, Bassham S, Cresko WA, Johnson EA. 2011. Local *de novo* assembly of RAD paired-end contigs using short sequencing reads. *PLoS ONE* 6:e18561.
- Feder JL, Egan SP, Nosil P. 2012. The genomics of speciation-with-gene-flow. *Trends in Genetics* 28: 342:350.
- Filzmoser P, Hron K, Templ M. 2018. *Applied compositional data analysis*. Springer International Publishing.
- Frost, SK, Epp LG, Robinson SJ. 1984. The pigmentary system of developing axolotls- A biochemical and structural analysis of chromatophores in wild-type axolotls. *Journal of Embryology and Experimental Morphology* 81:105–125.
- Fujii R. 1993. Cytophysiology of fish chromatophores. *International Review of Cytology* 143:191-255.
- Fujii R. 2000. The regulation of motile activity in fish chromatophores. *Pigment Cell Research* 13:300–319.
- Gehara M, Summers K, Brown JL. 2013. Population expansion, isolation and selection: novel insights on the evolution of color diversity in the strawberry poison frog. *Evolutionary Ecology* 27:297-824.
- Grether GF, Kolluru GR, Nersissian K. 2004. Individual colour patches as multicomponent signals. *Biological Reviews* 79:583–610.

- Grether GF. 2005. Environmental change, phenotypic plasticity, and genetic compensation. *The American Naturalist* 166:E115-E123.
- Greven H, Zanger K, Schwinger G. 1995. Mechanical properties of the skin of *Xenopus laevis* (Anura, Amphibia). *Journal of Morphology* 224:15–22.
- Hagemann S, Pröhl H. 2007. Mitochondrial paraphyly in a polymorphic poison frog species (Dendrobatidae; *D. pumilio*). *Molecular Phylogenetics and Evolution* 45:740-747.
- Haslam IS, Roubos EW, Mangoni ML, Yoshizato K, Vaudry H, et al. 2014. From frog integument to human skin: dermatological perspectives from frog skin biology. *Biological Reviews* 89:618-625.
- Hauswaldt JS, Ludewig AK, Vences M, Pröhl H. 2010. Widespread co-occurrence of divergent mitochondrial haplotype lineages in a Central American species of poison frog (*Oophaga pumilio*). *Journal of Biogeography* 38:711-726.
- Hoban S, Kelley JL, Lotterhos KE, Antolin M, Bradburd G, Lowry DB, et al. 2016. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist* 188:379-397.
- Hoekstra HE, Hirschmann RJ, Bunday RA, Insel PA, Crossland JP. 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313:101-104.
- Hoffman EA, Blouin MS. 2000. A review of colour and pattern polymorphism in anurans. *Biological Journal of the Linnean Society* 70:633-665.
- Karnovsky MJ. 1965. A formaldehyde-glutaraldehyde fixative of high osmolality for use in electron microscopy. *Journal of Cell Biology* 27:137-138A.
- Kikuchi DW, Seymoure BM, Pfenning DW. 2014. Mimicry's palette: widespread use of conserved pigments in aposematic signals of snakes. *Evolution & Development* 16:61-67.
- Kim J, Suh H, Kim S, Kim K, Ahn C et al. 2006. Identification and characteristics of the structural gene for the *Drosophila* eye colour mutant sepia, encoding PDA synthase, a member of the Omega class glutathione S-transferases. *Biochemical Journal* 398:451-460.
- Kindermann C, Hero JM. 2016. Pigment cell distribution in a rapid colour changing amphibian (*Litoria wilcoxii*). *Zoomorphology* 135:197-203.
- Knowles LL, Maddison WP. 2002. Statistical phylogeography. *Molecular Ecology* 11:2623-2635.
- Lavretsky P, DaCosta JM, Sorenson M, McGracken KV, Peters JL. 2019. ddRAD-seq data reveal significant genome-wide population structure and divergent genomic regions that distinguish the mallard and close relatives in North America. *Molecular Ecology* 28:2594-2609

- Lee YW, Gould BA, Stinchcombe JR. 2014. Identifying the genes underlying quantitative traits: a rationale for the QTN programme. *AoB Plants* 6:plu004.
- Lefcheck JS. 2016. piecewiseSEM: Piecewise structural equation modeling in R for ecology, evolution, and systematics. *Methods in Ecology and Evolution* 7:573-579.
- Mann ME, Cummings ME. 2008. Female preference for aposematic signal components in a polymorphic poison frog. *Evolution* 62:2334-2345.
- Maan ME, Cummings ME. 2009. Sexual dimorphism and directional sexual selection on aposematic signals in a poison frog. *PNAS* 106:19072-19077.
- Maddison DR, Maddison WP. 2000. *MacClade 4: Analysis of phylogeny and character evolution*. Sunderland, Massachusetts: Sinauer Associates.
- Maia R, Gruson H, Endler JA, White TE. 2019. pavo 2: new tools for the spectral and spatial analysis of colour in R. *Methods in Ecology and Evolution* 10:1097-1107.
- Malinsky M, Trucchi E, Lawson DL, Falush D. 2018. RADpainter and fineRADstructure: population inference from RADseq data. *Molecular Biology and Evolution* 35:1284-1290.
- Manceau M, Domingues VS, Linnen CR, Rosenblum EB, Hoekstra HE. 2010. Convergence in pigmentation at multiple levels: mutations, genes and function. *Philosophical Transactions of the Royal Society B* 265:2439-2450.
- McGraw KJ. 2006. Mechanics of carotenoid-based colouration. In *Bird Coloration, Volume I: Mechanisms and measurements*. Eds. Hill GE, McGraw KJ: 243-294. Harvard University Press, Cambridge.
- McRobie HR, Moncrief ND, Mundy NI. 2019. Multiple origins of melanism in two species of North American tree squirrel (*Sciurus*). *BMC Evolutionary Biology* 19:140.
- Montgomerie R. 2006. Analyzing colours. In *Bird Coloration, Volume I: Mechanisms and measurements*. Eds. Hill, GE & McGraw KJ: 90-147. Harvard University Press, Cambridge.
- Morrison, RL. 1995. A transmission electron microscopic (TEM) method for determining structural colors reflected by lizard iridophores. *Pigment Cell Research* 8:28-36.
- Müller F. 1879. *Ituna* and *Thyridia*: a remarkable case of mimicry in butterflies. *Proceedings of the Entomological Society of London* 20-29.
- Münkenmüller T, Lavergne S, Bzeznik B, Dray S, Thibaut J, et al. 2012. How to measure and test phylogenetic signal. *Methods in Ecology & Evolution* 3:734-756.
- Nakamura M. 2008. Sex determination in amphibians. *Seminars In Cell and Developmental Biology* 20:271-282.

- Nielsen HI, Dyck J. 1978. Adaptation of the tree frog, *Hyla cinerea*, to colored backgrounds, and the role of the three chromatophore types. *The Journal of Experimental Zoology* 205: 79-94.
- Nosil P, Funk DL, Ortiz-Barrientos D. 2009. Divergent selection and heterogeneous genomic divergence. *Molecular Ecology* 18:375-402.
- O’Leary SJ, Puritz JB, Willis SC, Hollenbeck CM, Portnoy DS. 2018. These aren’t the loci you’re looking for: principles of effective SNP filtering for molecular ecologists. *Molecular Ecology* 27:3193-3206.
- Paleo-Lopez R, Quintero-Galvis, JF, Solano-Iguaran JJ, Sanchez-Salazar, AM. A phylogenetic analysis of macroevolutionary patterns in fermentative yeasts. *Ecology and Evolution* 12:3851-3861.
- Pawlosky-Glahn V, Buccianti A. 2011. Compositional data analysis: theory and applications. John Wiley & Sons, Ltd.
- Peart CR, Dasmahapatra KK, Day JJ. 2018. Contrasting geographic structure in evolutionarily divergent Lake Tanganyika catfishes. *Ecology & Evolution* 8:2688-2697.
- Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2020. *nlme: Linear and Nonlinear Mixed Effects Models*. R package version 3.1-149
- Posso-Terranova A, Andres JA. 2017. Diversification and convergence of aposematic phenotypes: truncated receptors and cellular arrangements mediate rapid evolution of coloration in harlequin poison frogs. *Evolution* 71:2677-2692.
- Prager M, Andersson S. 2010. Convergent evolution of red carotenoid coloration in widowbirds and bishops (*Euplectes spp.*) *Evolution* 64:3609-3619.
- Prum RO, Torres R. 2003. Structural colouration of avian skin: Convergent evolution of coherently scattering dermal collagen arrays. *Journal of Experimental Biology* 6:2409–2429.
- Puritz JB, Matz MV, Toonen RJ, Weber JN, Bolnick DI, Bird CE. 2014. Demystifying the RAD fad. *Molecular Ecology* 23:5937-5942.
- R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rambaut AM, Suchard A, Xie D, Drummond. AJ. 2013. Tracer v.1.5.
- Rasotto MB, Cardellini P, Sala M. 1987. Karyotypes of five species of Dendrobatidae (Anura: Amphibia). *Herpetologica* 43:177-182.
- Revell LJ. 2012. phytools: An R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* 3:217-223.

- Reynolds, RG, Fitzpatrick BM. 2007. Assortative mating in poison-dart frogs based on an ecologically important trait. *Evolution* 61:2253–2259.
- Richards-Zawacki CL, Wang IJ, Summers K. 2012. Mate choice and the genetic basis for colour variation in a polymorphic dart frog: inferences from a wild pedigree. *Molecular Ecology* 21:3879-3892.
- Roberts JL, Brown JL, von May R, Arizabal W, Schulte R, Summers K. 2006. Genetic divergence and speciation in lowland and montane Peruvian poison frogs. *Molecular Phylogenetics and Evolution* 41:149-164.
- Rockman MV. 2012. The QTN program and the alleles that matter for evolution: all that's gold does not glitter. *Evolution* 66:1-17.
- Rodriguez A, Mundy NI, Ibanez R, Prohl H. 2020. Being red, blue and green: the genetic basis of coloration differences in the strawberry poison frog (*Oophaga pumilio*). *BMC Genomics* 21:301-317.
- Rogers RL, Zhou L, Chu C, Marquez R, Corl A, et al. 2018. Genomic takeover by transposable elements in the strawberry poison frog. *Molecular Biology and Evolution* 35:2913-2927.
- Rosenblum EB, Rompler H, Schoneberg T, Hoekstra HE. 2010. Molecular and functional basis of phenotypic convergence in white lizards at White Sands. *PNAS* 5:2113-2117.
- Rudh A, Rogell B, and Höglund J. 2007. Non-gradual variation in colour morphs of the strawberry poison frog *Dendrobates pumilio*: genetic and geographical isolation suggest a role for selection in maintaining polymorphism. *Molecular Ecology* 16:4284-4294.
- Santure AW, Poisson J, De Cauwer I, van Oers K, et al. 2015. Replicated analysis of the genetic architecture of quantitative traits in two wild great tit populations. *Molecular Ecology* 24:6148-6162.
- Saporito RA, Zuercher R, Roberts M, Gerow KG, Donnelly M.A 2007 Experimental evidence for aposematism in the dendrobatid poison frog *Oophaga pumilio*. *Copeia* 4:1006–1011.
- Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* 9:671-675.
- Schwander T, Libbrecht R, Keller L. 2014. Supergenes and complex phenotypes. *Current Biology* 24:R288-R294.
- Scrucca L, Fop M, Murphy TB, Raftery AE. 2016. mclust 5: clustering, classification and density estimation using Gaussian finite mixture models. *The R Journal* 8:289–317.
- Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, et al. 2010. Adaptation genomics: the next generation. *Trends in Ecology & Evolution* 25:705-713.

- Steiner CC, Weber JN, Hoekstra HE. 2007. Adaptive variation in beach mice produced by two interacting pigmentation genes. *PLoS Biology* 5:e219
- Stryjewski KF, Soreson MD. 2017. Mosaic genome evolution in a recent and rapid avian radiation. *Nature Ecology & Evolution* 1:1912-1922.
- Summers K, Cronin TW, Kennedy T. 2004. Cross-breeding of distinct color morphs of the strawberry poison frog (*Dendrobates pumilio*) from the Bocas del Toro archipelago, Panama. *Journal of Herpetology* 38:1-8.
- Summers K, Bermingham E, Weigt S, McCafferty S, Dahlstrom L. 1997. Phenotypic and genetic divergence in three species of dart-poison frogs with contrasting parental behavior. *Journal of Heredity* 88:8-13.
- Summers K, Cronin TW, Kennedy T. 2003. Variation in spectral reflectance among populations of *Dendrobates pumilio*, the strawberry poison frog, in the Bocas del Toro Archipelago, Panama. *Journal of Biogeography* 30:35-53.
- Swofford, D. L. 2002. Paup*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4.0b10. Sinauer Associates, Sunderland, MA.
- Szydlowski P, Madej JP, Mazurkiewicz-Kania M. 2015. Ultrastructure and distribution of chromatophores in the skin of the leopard gecko (*Eublepharis macularius*). *Acta Zoologica* 97:370-375.
- Taylor JD, Bagnara JT. 1972. Dermal chromatophores. *American Zoologist* 12:43-62.
- Taylor JD. 1966. Electron microscopy of iridophores in hypophysectomized *Rana pipiens* larvae. *American Zoologist* 6:587.
- Taylor JD. 1969. The effects of intermedin on the ultrastructure of amphibian iridophores. *General and Comparative Endocrinology* 12:405-416.
- Toews DPL, Hofmeister NR, Taylor SA. 2017. The evolution and genetics of carotenoid processing in animals. *Trends in Genetics* 33:171-183.
- Tuttle EM, Bergland AO, Korody ML, Brewer MS, Newhouse DJ, Mix P, Stager M, Betuel A, Cheviron ZA, Warren WC et al. 2016. Divergence and degradation of a sex chromosome-like supergene. *Current Biology* 26:344-350.
- Twomey E, Kain M, Claeys M, Summers K, Castroviejo-Fisher S, Van Bocxlaer IV. 2020. Mechanisms for color convergence in a mimetic radiation of poison frogs. *The American Naturalist*. 195:E132-E149.
- Twomey E, Johnson JD, Castroviejo-Fisher S, Van Bocxlaer IV. 2020b. A ketocarotenoid-based colour polymorphism in the Sira poison frog *Ranitomeya sirensis* indicates novel gene interactions underlying aposematic signal variation. *Molecular Ecology* 29:2004-2015.

- Twyman H, Prager M, Mundy NI, Andersson S. 2017. Expression of a carotenoid-modifying gene and evolution of red coloration in weaverbirds (Ploceidae). *Molecular Ecology* 27:449-458.
- van den Boogaart KG, Tolosana-Delgado R. 2013. *Analyzing compositional data with R*. Springer-Verlag Berlin Heidelberg
- Venables WN, Ripley BD. 2002. *Modern Applied Statistics with S*, Fourth edition. Springer, New York.
- Vestergaard JS, Twomey E, Larsen R, Summers K, Nielsen R. 2015. Number of genes controlling a quantitative trait in a hybrid zone of the aposematic frog *Ranitomeya imitator*. *Proceedings of the Royal Society, B* 282:20141950.
- Wang IJ, Shaffer HB. 2008. Rapid color evolution in an aposematic species: a phylogenetic analysis of color variation in the strikingly polymorphic strawberry poison-dart frog. *Evolution* 62:2742-2759.
- Wang IJ, Summers K. 2010. Genetic structure is correlated with phenotypic divergence rather than geographic isolation in the highly polymorphic strawberry poison-dart frog. *Molecular Ecology* 19:447-458.
- Weaver RJ, Santos ESA, Tucker AM, Wilson AE, Hill GE. 2016. Carotenoid metabolism strengthens the link between feather coloration and individual quality. *Nature Communications* 7:1-9.
- Weihs C, Ligges U, Luebke K, Raabe N. 2005. klaR Analyzing German Business Cycles. In Baier D, Decker R, Schmidt-Thieme L eds., *Data Analysis and Decision Support*, 335-343.
- Yang Y, Blumenkamp S, Dugas MB, Richards-Zawacki CL, Prohl H. 2019. Mate choice versus mate preference: Inferences about color-assortative mating differences between field and lab assays of poison frog behavior. *The American Naturalist* 193:598-607.
- Yang Y, Richards-Zawacki CL, Devar A, Dugas MB. 2016. Poison frog color morphs express assortative mate preferences in allopatry but not sympatry. *Evolution* 10:2778-2788.
- Yang Y, Richards-Zawacki CL. 2020. Male-male contest limits the expression of assortative mate preferences in a polymorphic poison frog. *Behavioral Ecology (in print)*
- Yeaman S, Whitlock MC. 2011. The genetic architecture of adaptation under migration-selection balance. *Evolution* 65:1897-1911.
- Zhang W, Westerman E, Nitzany E, Palmer S, Kronforst MR. 2017. Tracing the origin and evolution of supergene mimicry in butterflies. *Nature Communications* 8:12.